DOCUMENT RESUME

ED 264 541 CS 008 244

AUTHOR	Kamil, Michael L., Ed.; Moe, Alden J., Ed.
TITLE	Perspectives on Reading Research and Instruction. Twenty-Ninth Yearbook of The National Reading Conference.
INSTITUTION	National Reading Conference, Inc.
PUB DATE	80
NOTE	345p.; Proceedings of the Annual Meeting of the National Reading Conference (29th, New Orleans, LA, November 29-December 1, 1979).
PUB TYPE	Books (010) -- Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC14 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; Higher Education; Older Adults; Readability; Reading Achievement; Reading Attitudes; *Reading Comprehension; Reading Diagnosis; *Reading Instruction; Reading Research; Study Skills; *Teaching Methods
IDENTIFIERS	Orthography

ABSTRACT

Reflecting the increasing complexity and sophistication in reading research, the papers in this collection represent a wide spectrum of approaches, philosophies, viewpoints, and techniques of scholarly endeavor in their treatment of reading research and instruction. The 52 articles selected from those presented at the 1979 meeting of the National Reading Conference are arranged according to the following categories: (1) reading comprehension at the elementary and secondary school levels, (2) reading comprehension at the post-secondary level, (3) reading comprehension methodology, (4) review of research on comprehension, (5) reading and orthography, (6) instructional methods for reading, (7) reading and study skills, (8) attitudes and conceptions of reading, (9) reading and the aged, (10) assessment and reading, and (11) readability. The yearbook concludes with a list of all the papers presented at the conference. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

BEST COPY AVAILABLE

PERSPECTIVES ON READING RESEARCH AND INSTRUCTION

ED264541

Twenty-Ninth Yearbook
of
The National Reading Conference

MICHAEL L. KAMIL, Editor
University of Illinois at Chicago Circle
and
ALDEN J. MOE, Associate Editor
Purdue University

with the editorial
assistance of
CAROL A. DAVIS
and
LINDA K. GAUTHIER
Purdue University

Published by
The National Reading Conference, Inc.

1980

CS008 244

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Peter O. Allen

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Copyright, 1980
THE NATIONAL READING CONFERENCE, INC.
All rights reserved. This book or parts thereof
may not be reproduced in any form without
permission of the publishers

Distributed by
THE NATIONAL READING CONFERENCE
1230 17th Street, N.W.
Washington, DC 20036

The ideas expressed in the papers published
in the year book are those of the authors and do
not necessarily represent the viewpoints
of the National Reading Conference

Information concerning membership in The National Reading Conference and subscriptions to *Journal of Reading Behavior* and the *NRC Yearbook* may be obtained by writing to:

National Reading Conference
1230 17th Street, N.W.
Washington, DC 20036

OFFICERS AND BOARD OF DIRECTORS OF THE NATIONAL READING CONFERENCE, INC.

HARRY SINGER, *President*
University of California at Riverside

FRANK GREENE, *Vice-President*
McGill University

M. TRIKA SMITH-BURKE, *Treasurer*
New York University

JANET J. LARSEN, *Secretary*
University of Florida

J. JAAP TUINMAN, *Past President*
Simon Frasier University

THOMAS ESTES, *Board Member*
University of Virginia

THOMAS G. STICHT, *Board Member*
Human Resources Research Organization

IRENE ATHEY, *Board Member*
University of Rochester

KEN L. DULIN, *Board Member*
University of Wisconsin

MICHAEL McKENNA, *Convention Manager*
Wichita State University

MICHAEL L. KAMIL and ALDEN J. MOE, *Yearbook Editors*
University of Illinois at Chicago Circle and Purdue University

FRED W. OHNMACHT and RICHARD ALLINGTON, *Editors,*
Journal of Reading Behavior
State University of New York at Albany

PRISCILLA DRUM, *Publications Chairperson*
University of California

EDITORIAL ADVISORY COMMITTEE
NRC YEARBOOK

LINDA BAKER
University of Maryland

CAMILLE BLACHOWICZ
National College of Education

EMERY P. BLIESMER
Pennsylvania State University

CONNIE BRIDGE
University of Kentucky

ROBERT C. CALFEE
Stanford University

EUGENE H. CRAMER
University of Illinois at
Chicago Circle

PRISCILLA A. DRUM
University of California

GERALD G. DUFFY
Michigan State University

EDWARD FRY
Rutgers University

JANE HANSEN
University of New Hampshire

JEROME C. HARSTE
Indiana University

DALE D. JOHNSON
University of Wisconsin

ALBERT J. KINGSTON
University of Georgia

RON C. LESLIE
New York University

TERRY LOVELACE
University of Southwestern
Louisiana

MARY BETH MARR
State University of New York at
Albany

GEORGE H. McNINCH
University of Southern Mississippi

ALLAN R. NEILSEN
University of Calgary

EDWARD PARADIS
University of Wyoming

P. DAVID PEARSON
University of Illinois

CHARLES W. PETERS
Oakland Public Schools,
Pontiac, Michigan

SUSANNA PFLAUM
University of Illinois at
Chicago Circle

CHRISTOPHER J. RAMIG
Georgia State University

ALTON L. RAYGOR
University of Minnesota

LENORE RINGLER
New York University

HARRY SINGER
University of California, Riverside

DIXIE LEE SPIEGEL
University of North Carolina

ELIZABETH SULZBY
Northwestern University

ROBERT TIERNEY
University of Illinois

JOHN TOWNER
Western Washington University

ROBERT A. WISHER
Navy Personnel Research and
Developmental Center

BETTY H. YARBOROUGH
Old Dominion University

The Oscar S. Causey Award for Outstanding Achievement in College-Adult Reading

PRESENTATION OF THE OSCAR S. CAUSEY AWARD TO EARL F. RANKIN ON BEHALF OF THE NATIONAL READING CONFERENCE

1979

The Oscar S. Causey Award is made by the National Reading Conference to honor one of its esteemed members and to recognize an outstanding contribution to College-Adult Reading. This year's recipient is deserving on both counts.

He has been a National Reading Conference member since its earliest days and a friend and supporter for nearly every person in this room. He has been a Board member more than once. He has served as Vice-President, President and Past-President.

He has organized and run college reading centers; written research articles; constructed tests; reviewed books and tests; and, served as consulting editor to many journals, including our own *Yearbook* and *Journal of Reading Behavior*.

It is, therefore, with warmest thanks that the National Reading Conference *clozes* this presentation of the 1979 Oscar S. Causey Award to Earl F. Rankin.

The Oscar S. Causey Award for Outstanding Achievement in College-Adult Reading

PRESENTATION OF THE OSCAR S. CAUSEY AWARD TO
PAUL C. BERG
ON BEHALF OF THE NATIONAL READING CONFERENCE

1978

The Oscar S. Causey Award is made by the National Reading Conference to honor one of its esteemed members for outstanding contributions to Adult-College Reading. This year's Award winner is most deserving on both counts.

He has been a National Reading Conference member since its earliest days. He has edited the Yearbook. He has been a Board member on two different occasions. He has served as Vice-President, President and Past-President for the organization.

He has written over 50 articles and books. He has ably served as advisor to both State and Federal Governments (US and Foreign). He has served on the Board of Directors for other professional Associations in both Reading and in Speech and Hearing.

It is with warmest thanks that the National Reading Conference presents the 1978 Causey Award to Paul C. Berg.

TABLE OF CONTENTS

READING COMPREHENSION: ELEMENTARY AND SECONDARY

Selected Semantic Features of ESL Materials and Their Effect on Bilingual Students' Comprehension 1
T. Bean, T. Potter and C. Clark

An Analysis of Analogical Reasoning and Reading Comprehension Ability in Middle and Upper Grade Students 6
E. Cramer, E. Pascarella and C. Hendress

Preschoolers' Print Awareness: An In-depth Study of Three and Four Year Old Children 15
M. Hill

The Effects of an Induced Visual Imagery Strategy Upon the Reading Comprehension of Below Average Readers 22
P. Koskinen and L. Gambrell

Relation Between Levels of Prior Knowledge and the Organization of Recall 28
J. Langer

The Relationship Between Reading Performance and Awareness of Comprehension of Textbook Passages 34
J. Mize

The Effect of Passage Difficulty on Good and Poor Readers' Use of Author's Schema in Written Recall Protocols 38
R. Palmer, W. Slater and M. Graves

Sensitivity to Importance as a Predictor of Reading Comprehension 42
J. Pichert

Contextual Constraints and the Construct Validity of the Cloze Procedure 47
E. Rankin and S. Thomas

Strategies Children Use in Answering Questions 56
T. Raphael, P. Winograd and P. D. Pearson

Facilitating Children's Comprehension of Content Material 64
B. Taylor and S. Berkowitz

A Comparison of Subjects' Cloze Scores and Their Ability to Employ Macrostructure Operations in the Generation of Summaries 69
S. Thomas and C. Bridge

READING COMPREHENSION: POST-SECONDARY

Content and Cohesion in a Story Presented as Movie Versus Text 78
P. Baggett and A. Ehrenfeucht

Some Effects of Story Structure During Comprehension 88
R. Cirilo

Thematic Convergence in Reconstructed Passages in Mapping Activities 93
J. Davidson and D. Triplett

Comprehension: Of What the Reader Sees of What the Author Says 99
T. Estes and W. Shebilske

The Effect of Intersentential Relations on the Comprehension of Connected Discourse 105
V. Stone

READING COMPREHENSION: METHODOLOGY

Reading Comprehension Instruction in the Fourth Grade: Three Data-Gathering Methods 110
C. Hodges

Higher Level Features in Children's Stories: Rhetorical Structure and Conflict 117
C. Steinberg and B. Bruce

A Critical Look at Macro-Analyses of Text 126
R. Tierney and J. Mosenthal

REVIEW OF RESEARCH

A Short Survey of the Research on Text and Story Comprehension and Retention 131
L. Reder

READING AND ORTHOGRAPHY

Effects of Orthographic and Phonic Structures on Word Identification 154
C. Juel and R. Solso

Reading and Orthographic Structure 162
D. Massaro

The Effects of Selected Teacher Prompting Strategies on Oral Reading Performance 166
J. Niles

Learning to Spell in Spanish 172
C. Temple

INSTRUCTIONAL METHODS FOR READING

Using a Process Model as a Guide to College/Adult Reading Instruction 179
A. Forester

A Validated Procedure for Teaching Prefixes and Its Effect on Students' Ability to Assign Meaning to Novel Words 184
M. Graves and H. Hammond

The Effect of Repeated Reading of Taped Literature on Reading Achievement 189
J. Laffey, D. Kelly and B. Perry

The Effects of Three Instructional Activities on Children's Responses to the Open-Ended Question: What is a Word? 194
B. Pudis and E. Sulzby

The Effect of a Sequence Strategy on Teaching Reading Comprehension with the Cloze Procedure 201
E. Rankin, A. Haase, A. Stewart and E. Howard

READING AND STUDY SKILLS

Study Skills Mastery: Comparisons Between Teachers and Students on Selected Skills 207
E. Askov, K. Kamm, R. Klumb and M. J. Barnette

A Meta-Analysis of the Effect of Graphic Organizers on Learning from Text 213
D. Moore and J. Readence

A Comparison of Notetaking Strategies While Studying from Text 219
V. Orlando

ATTITUDES AND CONCEPTIONS OF READING

Changes in Reading-Related Attitudes During a One-Semester Pre-Service Course in Secondary Reading Methods 223
M. Dillingofski and K. Dulin

College Students' Metacognitive 226
V. Hare and C. Pulliam

The Influences of Grade and Pupil Ability Levels on Teachers' Conceptions of Reading 232
W. Metheny

Teacher Interns' Conceptualization of Reading Theory and Practice 238
R. Meyers and L. Ringler

READING AND THE AGED

An Evaluation of Literacy Programs for Mature Adults 243
J. Dinnan

Physical and Psychological Decrements Affecting Reading Among the Aged 245
M. Drotter

Reading Attitudes, Habits, Interests, and Motivations of the Elderly 250
K. Ngandu and B. O'Rourke

ASSESSMENT AND READING

Test Anxiety of College Students - Test of a Model 255
J. M. Bennett and D. Wark

The Need for Comparison Studies of College Students' Reading Gains in Developmental Reading Programs Using General and Specific Levels of Diagnosis 259
R. Flippo

A Passage-Dependent Test to Provide a Different View of Reading Comprehension 264
E. Fry and M. Reynolds

Statistical Models and Practical Significance in Reading Research 268
M. Hanson, W. Rupley, and J. McNamara

QASOR: A Framework for Qualitatively Analyzing Silent and Oral Reading 276
M. MacLean

Passage Independency: An Examination of Three Informal Reading Inventories Inventories 283
M. Marr and K. Lyon

The Effects of Age and Distractor Type on Maze Performance 288
M. McKenna and J. Miller

Literacy Achievement Profiles and Literacy Use - High and Low Achievers 292
R. Ruddell

Three Explanations of Reading Disability 301
T. Shanahan

READABILITY

Rauding and SEERing: Judging Global Readability 309
V. Froese

Text Structure and Comprehension: New Directions for Readability 313
J. Irwin

Categories and Strategies of Adaptation in Children's Reading Material 320
R. Kantor and A. Davison

Papers Presented at the 1978 meeting of the national Reading Conference 325

FOREWORD

Reading research has become increasingly an interdisciplinary activity. Specialists in linguistics, psychology, English, computer science, artificial intelligence and sociology, find reading a challenging field of inquiry. The articles in this volume represent nearly the full range of interests in reading research and instruction. We believe that the information in these articles moves us, in some way, towards a better understanding of the work of some of our colleagues in "other" disciplines. A consequence will be improved instructional practices. We feel the membership of the National Reading Conference is vitally working to those ends.

The quality of this productivity is evidenced by the fact that this volume contains only 40% of the total number of manuscripts submitted. To all those who submitted manuscripts we offer a sincere thanks. To those members who served as the Editorial Advisory Committee we offer a special thanks for their critical reviews and constructive comments. And to those authors who responded so quickly with revised and complete manuscripts, we offer a very special thanks.

To Carol Davis and Linda Gauthier we extend special appreciation for their diligence and editorial expertise. They made the arduous task of producing this volume much easier. To the staff at Printing Services at Purdue we extend our full appreciation for their work in composing and printing this volume. And finally, as this foreword is being written and as the last of the galley proofs are sent out, one of the editors works around packing crates ready to move to a new academic setting — such is the nature of editors and educators.

Michael L. Kamil
Alden J. Moe

THOMAS W. BEAN
TOMAS C. POTTER
California State University

CHARLES H. CLARK
Western Illinois University

SELECTED SEMANTIC FEATURES OF ESL MATERIALS AND THEIR EFFECT ON BILLINGUAL STUDENTS' COMPREHENSION

Prose comprehension involves a complex interplay of reader and text variables. Two semantic features of prose that continue to interest researchers are story structure and pronoun reference. The present study explored the effect of story structure and pronoun reference on the comprehension of middle grade students who were speakers of English as a second language (ESL).

Previous studies of childrens' story comprehension suggest that the level of information recalled by a reader will reflect the degree to which a selection unfolds as a predictable problem-solving structure (Rumelhart & Ortony, 1977). A number of researchers have devised systems for analyzing the hierarchical structure of stories (Rumelhart, 1977; Mandler, 1978; Stein & Glenn, 1978). These story grammars represent ideal problem-solving scenarios that should assist children in comprehending the "gist" of a narrative. Studies exploring story recall indicate that readers display a tendency to remember information that is prominent in the story structure while distorting or deleting more trivial details (Rumelhart, 1977).

A second semantic feature that appears to effect middle grade childrens' comprehension is anaphora, particularly pronoun reference. Marshall (1979) suggested that distant pronoun antecedents disrupt textual cohesion making comprehension difficult for the reader. Indeed, Gottsdanker (1975) found that middle grade students obtained higher comprehension scores on passages containing low rather than high frequencies of pronouns.

If story structure and pronoun reference play a role in the prose comprehension of native English speakers, how are these features treated in narrative selections intended for ESL students?

A representative story entitled "The Chase Twins" from the *Miami Linguistic Readers* (D. C. Heath, 1966) was analyzed for use in the present study. This particular reading series is intended for beginning and middle grade students whose native language is Spanish.

The plot of the target story concerned twins who planned to find a stolen ruby ring. The twins' father was one of the policemen responsible for catching the robber in the story. As the plot unfolds at the breakfast table, the twins' father unknowingly piques their interest in solving the mystery of the missing ring by discussing various clues with them. This opening portion of "The Chase Twins" was selected for use in the present study since the complete story was considered to be too lengthy for experimental purposes. This section of the story contained the first three categories of Stein and Glenn's (1978) story grammar: setting, initiating event, and goal formation. However, the movement of the story was continually disrupted by the introduction of trivial events such as an argument the twins were having about what to do that day.

In addition to some observable problems with story structure, the selection displayed the artifacts of traditional readability formulas. A large number of easy to pronounce sight words, particularly pronouns with obscure and sometimes distant referents were noted. In some cases, the reader was required to link a pronoun with an antecedent that was introduced over 18 steps back in the story.

A pilot study was conducted in which fourth grade ESL subjects were asked to read "The Chase Twins" silently and retell it in their own words. Their taped retellings revealed a tendency to recall trivial information including the twins' argument while completely missing the goal formation category of finding the ring. Moreover, subjects in the pilot study seemed to experience moderate difficulty identifying the anaphoric reference that linked their father with the other policemen. Thus, it appeared that these two semantic features, story structure and pronoun reference, were related to ESL students' understanding of this selection.

The present study was conducted to determine the degree to which a predictable story structure and a reduction in obscure pronoun referents would influence ESL students' comprehension of this story.

METHOD

Subjects

The subjects for this study were 45 fourth and fifth grade students whose native language was Spanish. They were enrolled in a public elementary school in a suburb of Los Angeles. All subjects were billingual in Spanish and English according to oral fluency scores on the *Bilingual Inventory of National Language* (Herbert, 1975).

Materials

Three story conditions were employed in this study. A portion of a representative ESL selection constituted the first story condition. In contained 583 words and was judged to be at third grade difficulty on the Raygor readability graph (Raygor, 1977).

The second story condition involved simply rewriting sentences in such a way as to make obscure pronoun referents more explicit without modifying the original story structure. Twenty such pronouns were identified and subsequently rewritten. Thorndyke's (1977, p. 82) definition of a proposition as "a clause or sentence containing an action or stative verb" was adopted for this procedure. For example, in the development section of the original story, the single proposition and related pronoun, "*it* really is a twin mystery," was rewritten as two propositions, i.e. "*the stolen ruby ring* is a twin mystery." This resulted in a selection that was 624 words long with Raygor readability of fourth grade.

The third story condition involved the deletion of trivial events in the original selection that tended to detract from a predictable problem-solving structure. For example, the setting contained 13 propositions in the original condition. This was reduced to just two essential propositions in the third story condition. Pronoun referents were made explicit as in the second story condition. These modifications to the original story resulted in a shorter selection of 427 words with a higher readability rating of fifth grade on the Raygor.

Procedure

The 45 fourth and fifth grade subjects were randomly assigned to one of three story conditions. This resulted in 14 subjects reading the original story, 16 the explicit pronoun version, and 15 the predictable story structure/explicit pronoun condition.

The subjects met individually with an examiner. After establishing rapport, subjects were advised to read the story silently so they could retell it in their own words. Following their reading of the story, subjects' retellings were audio taped for later protocol analysis. Subjects then responded orally to a 10 item probed recall containing wh-questions. Only the probed recall data is reported in this paper.

The probed recall was designed to assess subjects' understanding of the overall "gist" of the story (i.e. find the missing ring and solve the mystery), as well as their ability to identify obscure pronoun referents. For example, question nine, "What do the twins decide to do?" (play dectective and find the ring), related to the goal or "gist" of the story. Question six, "Why does Mr. Chase know so much about the robbery?" (he is one of the policemen who caught the robber), revealed the degree to which subjects were able to resolve an obscure pronoun referent. The probed recalls were audio taped and scored by two raters yielding an inter-rater reliability of .96 based on a Pearson r for the paired scores.

RESULTS

The mean probed recall scores for each of the three story conditions were: 2.64 on the original; 3.18 on the pronoun rewrite; 5.06 on the story structure/pronoun rewrite. An analysis of variance indicated a significant difference in comprehension for the three story conditions ($F(2.42) = 4.68, p < .025$). Scheffé comparisons of the group means revealed a significant difference for the story structure/pronoun rewrite condition and the original version (difference = 2.42, $p < .05$). Comparison of the pronoun rewrite condition and the original version did not reach significance (difference $<$ .76).

DISCUSSION

Although pronouns may increase sentence level processing demands for a reader, recent research has begun to explore anaphora in the context of continuous discourse. In some cases, the results of these studies imply that pronoun resolution is much less of a problem for middle grade children than previously believed. For example, in the present study ESL subjects' comprehension was not significantly improved by making obscure pronoun referents explicit. This finding is consistent with a recent, developmental study of pronoun resolution conducted by Barnitz (1980). He found that fourth graders were able to make the bridging inferences required for pronoun resolution in well-formed narratives. Barnitz reasoned that a child's prior knowledge and expectations concerning stories overshadowed any problems with sentence level cohesion. The findings of the present study suggest that ESL subjects processed "The Chase Twins" in a similar, top-down fashion.

Middle grade ESL students were able to comprehend the "gist" or goal of "The Chase Twins" significantly better when it was restructured to conform to a predictable, problem-solving scenario. Rumelhart (1977) and Bower (1978) found that readers typically identified with the protagonist's pursuit of a goal in a story. More importantly, they used this feature to bind together the events of a story in memory. In contrast, a story setting is usually of minor importance to the reader's recall of a story (Rumelhart, 1977). Yet in the original version of "The Chase Twins," there were 13 propositions in the setting which centered on the twins' argument about what to play that day — cops and robbers, or hospital. This argument, comprising a familiar and salient event for children may have represented a memorable sub-goal that detracted from their recall of the "gist". This argument portion of the setting was deleted in the restructured version which undoubtedly contributed to the subjects' significantly greater comprehension of the story.

Subjects achieved the highest mean comprehension scores on the most difficult story condition in terms of a traditional readability rating (i.e. fifth grade). The original story was the least difficult yet subjects achieved the lowest mean probed recall scores on this version. This finding suggests that narratives for ESL students that are constructed according to the traditional tenets of readability may be easy to recode orally, yet difficult to comprehend.

The present study explored childrens' comprehension of a single story designed for ESL students. While the use of one story alone may be viewed as a limitation restricting generalizability, the findings are consistent with a growing body of research in story structure and anaphora. Carver (1978) recommended replication as a means of strengthening inferences that extend beyond the particular materials used in an experiment. Replication of the present study with a population of native English speakers would broaden its generalizability. Indeed, studies exploring the interacion of macro and microlevel features of prose should serve to increase our understanding of the reading process with respect to a variety of learners.

REFERENCES

Barnitz, J. G. Syntactic effects on the reading comprehension of pronoun-referent structures by children in grades two, four, and six. *Reading Research Quarterly*, 1980, *15*, 268-289.

Bower, G. H. Experiments on story comprehension and recall. *Discourse Processes*, 1978, *1*, 211-231.

Carver, R. P. Sense and nonsense about generalizing to a language population. *Journal of Reading Behavior*, 1978, *10*, 25-33.

Gottsdanker, A. E. *The effect of some anaphoric structures on reading comprehension.* Unpublished masters thesis, University of California at Santa Barbara, 1975.

Herbert, C. H. *Bilingual Inventory of National Language.* San Bernardino, California: CHEC Point Systems, 1975.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes, 1978, 1*, 14-35.

Marshall, N. Readability and comprehendability. *Journal of Reading*, 1979, *22*, 542-544.
Miami Linguistic Readers. Lexington, Mass.: D. C. Heath, 1966.
Raygor, A. L. The Raygor Readability Estimate: A quick and easy way to determine difficulty. In P. D. Pearson (Ed.), *Reading: Theory, research, and practice*. Twenty-sixth Yearbook of the National Reading Conference. Clemson S.C.: The National Reading Conference, 1977.
Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.
Rumelhart, D. E. & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.
Stein, N. L. & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *Advances in discourse processes: New directions*. Norwood, New Jersey: Ablex, 1978.
Thorndyke, P. W. Cognitive structures in comprehension and memory. *Cognitive Psychology*, 1977, *9*, 77-110.

EUGENE H. CRAMER
ERNEST T. PASCARELLA
University of Illinois
at Chicago Circle

CHERYL HENDRESS
Educational Service Region
Bureau County, Illinois

AN ANALYSIS OF ANALOGICAL REASONING AND READING COMPREHENSION ABILITY IN MIDDLE AND UPPER GRADE STUDENTS

PURPOSE

The purpose of this study was to analyze certain relationships between reading comprehension and analogical reasoning. By performing a number of comparisons of pupils' scores on a standardized reading achievement test with their scores on a specially constructed verbal analogies test, the study sought at least a partial answer to the question: "If a child can decode easy words, can that child reason with those words?"

In the more than sixty years since Thorndike concluded that reading is reasoning (Thorndike, 1917), various explanations of the reading-reasoning process have been advanced. In recent years, a number of researchers and theorists have turned away from earlier behavioral explanations and moved toward cognitive explanations of the reading comprehension process. A current line of investigation involves schema theory, and the writing of two of its exponents provided partial impetus for this study. Rumelhart and Ortony (1977) have suggested that in schema theory the entire comprehension process may be regarded in terms of analogical reasoning:

> Thus, the reading strategies which people normally employ can readily be incorporated into a schematic representation of knowledge such as the one we propose. These principles may not only include, but may also transcend, the "laws of thought" which are described in textbooks. Finally, at a more general level, it is interesting to note that one can regard the entire comprehension process in schema theory as itself being a case of analogical reasoning. When we determine that a situation fits a certain schema, we are in a sense determining that the current situation is analogous to those situations from which the schema was originally derived. Moreover, when we make inferences about unobserved aspects of the situations we are, in effect, assuming their existence by *analogy* from the situations from which the schemata were derived.

Thus, this study proceeded from the assumption that the ability to solve verbal analogy problems involves a complex network of associations derived from past experiences and incorporated into a child's schema.

The mental manipulations by which analogy problems are processed and solved have been investigated by Sternberg (1977, 1979) and Whitely

(1977). Sternberg argues for developmental patterns in the ability to recognize higher-order analogical relations. Whitely has derived discrete categories of verbal analogies from investigations of college students' performance on an analogies classification task. Neither studied verbal analogies and reading.

A common complaint from classroom teachers is that while many children seem able to decode easy words, they seem less able to understand prose passages in which these easy words appear. What was attempted in this study was the construction of a valid and reliable verbal analogies test, using only easy words, that would be appropriate for use with children in the middle and upper elementary grades. If found to be reliable, the verbal analogies test could be used to examine certain relationships between reading comprehension and analogical reasoning, and ultimately to become a tool for classroom teachers to use in determining a child's ability to reason analogically with the words s/he is able to decode.

METHOD

Instruments

Verbal Analogies Test — Form 2 (VA-2). Students' analogical reasoning ability was assessed with a specially constructed test, VA-2. Using only words taken from the "Dale List of 769 Easy Words," the test items were assumed to be appropriate for any student reading at the third grade level or above. Following the work of Whitely (1977) and Hafner (1977), nine categories of verbal analogies were derived. The nine categories are named, defined, and exemplified in Figure 1. Six test items were constructed from each category, or 54 items in all. Aptness of test items to their assigned categories was verified by a panel of twelve graduate students in reading, with an interjudge reliability coefficient of .88.

Each analogy problem in the test followed the pattern *A* is to *B*, as *C* is to ____. A study of Ace and Dawis (1976) concluded that blank position in verbal analogy items is probably not a significant factor influencing the difficulty of items. The order of analogy categories was selected randomly for the first nine items; then, the items were rotated in the same order for the balance of the test. Location of correct answers among the multiple choice distractors was also selected randomly.

The VA-2 Test was administered by one of the authors to the subjects on April 4, 5, and 9, 1979. Subjects were permitted as much time as necessary to complete the test. Beginning and ending times were noted for all groups, with total test times ranging from 11 to 24 minutes.

SRA Achievement Tests, Level E — Form 1. Students' reading ability was assessed with the Vocabulary and Reading Comprehension subtests of the Science Research Associates Achievement Series, Level E — Form 1, 1978. The SRA Achievement tests were administered by local school personnel on April 23 and 24, 1979.

Subjects

Subjects in this study were 307 students in grades four through eight from 12 randomly selected classroom groups in four schools in an Illinois town of

FIGURE 1.
Verbal Analogy Categories, Definitions, and Examples

Category	Definition	Example
Functional	a group of relationships in which one object serves some activity or function for the other	Fish is to Water as Bird is to ________
Part/Whole	a group of relationships in which one object is a part of another, or contains the other	Toe is to Foot as Finger is to ________
Contrasts/ Opposites	a group of relationships in which the terms have opposite or contrasting meanings	Before is to After as Early is to ________
Quantitative	a category in which the relationship between terms is size, order, or number	ABA is to BBA as 121 is to ________
Word Pattern	a group of terms in which the relationship is based upon non-semantic properties of the words	Win is to An as Wind is to ________
Conversion/ Predication	a group of relationships in which one object can be converted into the other by some natural or artificial process; or in which one term is assumed to be a natural or commonly associated predication of the other	Story is to Tell as Book is to ________
Class Naming	a class relationship in which one object belongs to a class named by the other	Red is to Color as Sweet is to ________
Class Membership	a category of classes in which both terms are examples of some class which must be educed	Red is to Green as Blue is to ________
Similarities/ Synonyms	a category of similar or synonymous word meanings	Fine is to Nice as Large is to ________

8000, located about 120 miles from Chicago. The town is a stable community of middle-class merchants, farmers, and employees and managers engaged in light industry. Students ranged from low-normal to high-normal in ability. A few students who had been labeled as "learning disabled" were excluded from this study.

Questions Addressed in the Study

Six major questions were addressed in this study; the first three dealt with the VA-2 instruments, and the last three dealt with relationships among analogical reasoning and reading achievement:

1. Is the Verbal Analogies-Form Two a reliable test?
2. Is the division of analogical reasoning test items into nine categories a justifiable one?
3. Is the VA-2 Test a useful instrument for predicting reading achievement?
4. Is there a relationship between subjects' ability to reason analogically and their grade level or sex?
5. Is there a relationship between subjects' scores on each of the nine verbal analogy categories and their reading achievement scores?
6. Is there a relationship between subjects' grade levels and their scores on each of the nine verbal analogy categories?

RESULTS

The first question of the study concerned the internal consistency reliability of the VA-2 instrument. The alpha (internal consistency) reliability of the 54 item score across the 307 subjects was computed using the reliability program of the Statistical Package for the Social Sciences. The computed reliability coefficient of .86 was judged to be high enough to warrant further analyses with the instrument.

The second question of the study concerned the utility of the nine verbal analogies categories. Our approach was to determine differences among the categories with respect to average difficulty and to determine the range and average covariation among the categories. The overall mean scores (as well as the standard deviations) of the nine VA-2 categories are shown in ascending order of difficulty in Table 1 (variables 1-9). As indicated, the Contrasts/Opposites category was the least difficult for all students while Word Pattern was the most difficult.

The intercorrelations among the nine VA-2 categories are also shown in Table 1. The correlations ranged from a low of .21 (between Contrasts/Opposites and Class Naming) to a high of .51 (between Part-Whole and Word Pattern). The median correlation was .37. Thus, while there was modest covariation among the nine categories, ranging from about 4 percent to 26 percent and averaging about 14 percent, no one category shared more than 26 percent of its variance with any other. This suggests reasonable independence among the nine VA-2 categories.

TABLE 1

Means, Standard Deviations and Intercorrelations of Nine Verbal Analogies Scales (1-9) and Three Measures of Reading Achievement (10-12)[a]

Variable		2	3	4	5	6	7	8	9	10	11	12	Mean	SD
1	Contrasts/Opposites	40	28	35	21	36	34	30	37	36	41**	41**	5.57	.76
2	Quantitative		26	36	26	41	50	36	48	49**	51**	54**	4.83	1.03
3	Similarities/Synonyms			27	30	42	37	32	42	24	30	30	4.77	1.12
4	Class Membership				37	41	43	41	48	38	40	42	4.44	1.28
5	Class Naming					41	36	36	41	27	29	31	4.40	.93
6	Part-Whole						49	50	51	43	48*	50*	4.08	1.18
7	Functional							46	50	50**	56**	58**	4.07	1.29
8	Conversion/Prediction								49	41	48**	49*	3.55	1.52
9	Word Pattern									45	49	50	3.40	1.73
10	Vocabulary Grade Equivalent										79	93	7.63	2.64
11	Comprehension Grade Equivalent											94	8.08	3.03
12	Reading Total Grade Equivalent												7.76	2.70

[a] Decimals omitted for correlations

*Indicates that regression analysis beta weight was significant at $p \leqslant .05$

**Indicates that regression analysis beta weight was significant at $p \leqslant .01$

A third question of the study was the utility of the scales in predicting reading achievement. Significant associations were found between subjects' VA-2 total scores and their scores on three SRA Reading Achievement subtests. The zero order correlations, all significant at $p < .01$, were as follows: vocabulary .58 ($r^2 = .34$), comprehension .65 ($r^2 = .42$), and reading total .67 ($r^2 = .45$). When each of the three reading achievement measures was regressed on VA-2 total score, holding student sex and grade level constant, the increase in explained variance (R^2) due to VA-2 total was as follows: vocabulary .18, comprehension .24, and total reading .23. Each R^2 increase was significant at $p < .01$.

In order to investigate the relationships between the VA-2 scores and student sex and grade level, which was the fourth question posed in the study, a 2 (sex) by 5 (grade level) least squares analysis of variance was conducted on the VA-2 total score. The main effect for sex was non-significant ($F = 0.68$, df = 1/297, $p < .40$), but there was a significant grade level effect ($F = 14.59$, df = 4/297, $p < .001$). The sex X grade level interaction was non-significant ($F = 1.04$, df = 4/297, $p < .35$). With sex held constant, the adjusted VA-2 total score means for grade level were as follows: grade 4 = 36.06, grade 5 = 36.58, grade 6 = 40.52, grade 7 = 41.56, grade 8 = 43.80.

The fifth question of the study dealt with relationships among students' scores on each of the verbal analogy categories and their reading achievement scores. Accordingly, the three reading achievement subtests were regressed on the nine individual VA-2 categories. The total variance (R^2) explained in the vocabulary, comprehension, and reading total scores by the set of verbal analogies subscales was 38%, 46%, and 50%, respectively (each significant at $p < .01$). Standardized partial regression coefficients (or beta weights) were used to determine the relative contribution of each of the nine analogy categories to the prediction of the reading achievement measures with all other categories held constant. Of all nine subscales, only two, Quantitative analogies and Functional analogies, had positive and significant ($p < .01$) beta weights with all three reading achievement measures.

The sixth question of the study concerned relationships between subjects' grade levels and sex and their scores on each of the nine verbal analogy categories. Two discriminant analyses were conducted to determine the extent to which scores on the nine individual verbal analogies subscales were related to student sex and grade level. The multivariate F for sex was only 0.96 with 9/297 df, $p < .46$. However, the multivariate test for grade level was significant at $p < .001$ ($F = 3.12$ with 36/1103 df). Scores on three of the nine analogy subscales, Quantitative, Part—Whole, and Functional analogies, had significant ($p = .01$), positive relationships with grade level, with all other scales held constant.

Furthermore, a plot of the standardized group centroids revealed three rather distinct clusters of grade-related performance across the nine analogies categories. (A minus value is comparatively low performance, while a positive value signifies a relatively higher performance.) Grades 4 and 5 tended to cluster together with centroid values of -.63 and -.43 respectively; grades 6 and 7 clustered together with centroids of .35 and .24 respectively; and grade 8 was a high performance outlier with a centroid of 1.01. Tests for centroid differences indicated non-significant differences between grades 4 and 5 and between grades 6 and 7. However, both grades

4 and 5 were significantly lower than either grade 6 or 7, and grade 8 was significantly higher than all other grades.

SUMMARY AND DISCUSSION

Reliability of the VA-2 Test

The Verbal Analogies — Form 2 Test, with a split-half reliability coefficient of .86, may be regarded as sufficiently reliable for use with pupils in grades four through eight who are reading in the low-normal to high-normal range. In a revision which is planned for the test, specific test items found to be contributory to a decrease in overall reliability will be eliminated or rewritten.

Justification of the Nine-Categories Construct

Intercorrelations among the nine categories of verbal analogies seemed high enough (.21 to .51) to suggest that they may be tapping a common analogical reasoning factor. On the other hand, the intercorrelations were not so high as to suggest that the individual scales are not also assessing unique dimensions of analogical reasoning. No more than 26% of the variance of any one scale was shared with any other single scale.

VA-2 Test as a Predictor of Reading Achievement

A tentative conclusion of this study is that the VA-2 Test appears to be a significant predictor of reading achievement, even with student grade level and sex held constant. This conclusion is based upon the findings discussed below:

Relationships between Analogical Reasoning and Sex, Grade Level, and Reading Achievement

Analogical reasoning ability, as measured by the VA-2 Test, correlated significantly with vocabulary, comprehension, and total reading scores as measured by a standardized reading achievement test. The theoretical speculation of Rumelhart and Ortony (1977), cited earlier in this report, would seem to gain credence from the findings of this study: a substantial amount of the variance in reading comprehension scores is explained by the analogical reasoning variable.

A related finding is that subjects' analogical reasoning ability appeared to increase with grade level. The implications of this finding are many and complex. It may be that a developmental factor may be operating, as propounded by Spearman (1977) and earlier by Piaget (1950); or that as a child's schema becomes more rich and complex from increased years of experience with his world, analogical reasoning may become a more powerful strategy for dealing with the unknown. These are, however, highly speculative notions which certainly seem to warrant further investigation.

There was no apparent relationship between sex of subjects and their analogical reasoning ability.

Relationships of the Nine Verbal Analogy Categories with Reading Achievement

Two of the nine verbal analogy categories, Quantitative and Functional, emerged as most powerful in their ability to predict reading achievement. Conversely, the Synoym/Similarities and the Class Naming categories of verbal analogies made little contribution to the prediction of reading achievement. The other five categories, Contrasts/Opposites, Class Memberships, Part-Whole, Conversion/Prediction, and Word Pattern, made modest prediction contributions. If, after replicative study, the relative amounts of contribution to prediction remain reasonably constant, classroom teachers may be encouraged to concentrate instruction in analogical reasoning accordingly.

Grade Level Differences in Analogical Reasoning Categories

Three of the nine verbal analogy categories, Quantitative, Part-Whole, and Functional, contributed significantly to grade level discrimination; that is, subjects at different grade levels performed quite differently on analogy test items in these three categories. When this performance differential was further examined, it was found that subjects' scores tended to cluster. Fourth and fifth grade subjects scores clustered at the lowest level (centroids for grades 4 and 5 were -.63 and -.43 respectively); sixth and seventh grade subjects' scores clustered at a middle level (centroids for grades 6 and 7 were. 35 and .24 respectively); and eighth grade subjects' scores stood alone at the highest level (centroid for grade 8 was 1.01). This cluster progression suggests that there may be a maturational factor involved in students' ability to solve verbal analogy problems, particularly in the Quantitative, Part-Whole, and Functional Categories.

Further study will need to be done to determine implications of this finding. It may be that these categories should be the ones in which classroom teaching materials need to be developed. Particular attention should be given to Quantitative and Functional analogies, since these two verbal analogies also made significant contributions to prediction of reading achievement.

LIMITATIONS OF THE STUDY

There are at least two major limitations of this study which need to be considered in any analysis of the findings reported. First, since no IQ scores could be obtained for the subjects in the study, it was not possible to control for the possible effects of intelligence on the analogical reasoning or reading achievement scores. A replication study is planned in which IQ scores will be available and the influence of this important variable can be taken into account.

A second limitation of the study was that no attempt was made to control for subjects' differences in background knowledge. Although all the subjects came from a relatively stable community and had similar backgrounds, it is likely that there were considerable differences in their experiential knowledge and, thus, differences in their familiarity with concepts being used in the analogy items. Of course, it could be said that this is a weakness of nearly every test, but it seems especially critical in a test of reasoning ability. Further studies of reading-reasoning relationships need to attempt to control this factor even beyond our attempt to do so with a list of "easy" words.

SUMMARY CONCLUSION

A final statement is at least a partial answer to the question posed at the beginning of this report: "If a child can decode easy words, can that child reason with those words?"

The answer would seem to be, "Not necessarily." All words used in the VA-2 Test were "easy" words, selected from the Dale List of 769 Easy Words. No decoding difficulties were noted or reported for any of the 307 subjects, all of whom are reading at or above grade level. However, not all children seemed able to reason analogically with these same easy words. In general, the VA-2 Test increased in efficiency as a predictor of reading achievement as students' grade levels increased. Whether this is strictly a maturational factor, or whether it is possible to increase students' reading achievement by increasing their analogical reasoning ability through instruction remains open for further study. If, as seems fairly certain, a child's ability to read is closely associated with the child's ability to reason, it seems reasonable that some direct instruction in reasoning by analogy is in order.

REFERENCES

Ace, M. C. & Dawis, R. V. Item structure as a determinant of item difficulty in verbal analogies. *Educational and Psychological Measurement*, 1976, *33*, 143-149.

Hafner, L. E. *Developmental reading in middle and secondary schools*. New York: Macmillan, 1977.

Piaget, J. *The psychology of intelligence*. New York: Harcourt, Brace, 1950.

Rumelhart, D. E. & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisiton of knowledge*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Sternberg, R. J. *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Sternberg, R. J. The nature of mental abilities. *American Psychologist*, 1979, *34*, 214-230.

Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, 1917, *8*, 323-332.

Whitely, S. E. Relationships in analogy items: A semantic component of a psychometric task. *Educational and Psychological Measurement*, 1977, *37*, 725-739.

MARY W. HILL
Westminster College, Pa.

PRESCHOOLERS' PRINT AWARENESS: AN IN-DEPTH STUDY OF THREE AND FOUR YEAR OLD CHILDREN

Based on the assumption that children need to be taught the mechanical skill of writing, writing instruction in many first-grade classrooms primarily consists of children learning the form of written language. First-graders are taught how to write the graphic symbols, to write left to right and top to bottom, and to match sound symbols with letter symbols. As these skills are taught children are also taught to spell simple words and then to use those words in their own stories, usually called creative stories. Typically, initial instruction includes: (1) children overwriting, tracing over, either published or teacher made materials, (2) children under writing, writing underneath, either published or teacher made materials and/or (3) children copying their own previously dictated language experience stories which have been scribed by the teacher.

Instruction which is predicated upon the above assumption places emphasis on the form rather than on the function of written language. That is, writing is equated with forming the graphic representation. Little emphasis is placed on writing as a process in which one conveys a message and on writing having different functions. There is limited opportunity for the writer to focus on the text as the integral part of the writing process. What the young writer intuitively knows about language, how it functions and its purpose, seldom becomes evident when the instructional focus is on the product rather than on the process. It follows then that evaluation of young children's writing will be dependent upon their performance rather than upon their competence and performance.

When emphasis is on performance, or form, an assumption has been made that children have limited, if any, competence in writing. Thus children are considered to have little or no awareness of print and how it functions. Recent research (C. Chomsky, 1970; Clay, 1975; Harste & Burke, 1978; Harste, Burke, & Woodward, 1979; Read, 1975) however suggests that children know a great deal about written language at an early age. This study then attempts to examine the writing process of a group of children who are in nursery school and kindergarten and have had no formal writing instruction prior to the last portion of their kindergarten year.

Survey of the Study

From a socio-psycholinguistic perspective three assumptions are being made by this researcher concerning preschoolers' print awareness. First, young children are aware of written language and systematically evidence a control of representational print; second, young children are aware that the purpose of written language is to communicate; and third, young writers have a schema for written language in that they expect it to have different functions.

Description of Data Collection

This study was begun with a random selection of six children attending an early childhood program in an elementary school. Due to the age of the children no effort was made to select the children other than their willingness to participate in a one-to-one writing activity with this researcher. The study began with three children, two girls and a boy, from the three-year-old room and three children, two girls and a boy, from the four-year-old room. Since two of the children moved at the end of the first year of the study, the analysis of the data is made on the following children: Erica (3 yr., 2 mo.), Nicole (3 yr., 10 mo), Alison (4 yr., 5 mo.), and Megan (4 yr., 7 mo.).

The linguistic environment for the three-year-olds' room and the four-year-olds' room was similar. It consisted of the typical classroom messages and signs; i.e. color and number words, safety messages (STOP, GO, BE CAREFUL), class rosters (names on lockers, room responsibilities). Books and writing materials, paper, pencils, crayons, and paints, were available for the children to use. The children went to the school library on a regular schedule and the teacher, or the assistant teacher, read to them frequently.

At the beginning of the five-year-olds' program the linguistic environment did not differ except there were more books available and the teacher read to the children as a planned activity each day. By the end of the year planned writing activities were introduced which included children being taught to form letters, identify them, and relate letter names and sounds. There was no emphasis on children writing by themselves. Instead the children dictated and the teacher acted as scribe. They then copied what had been written for them by the teacher.

Procedure

Data were collected on a periodic basis during the academic years 1977-1978 and 1978-1979. The researcher and each child sat at a table on which there were pencils, ball point pens, crayons, and blank writing paper. The verbal instruction given to each child at the initial meeting was as follows:

> I would like for you to draw a picture for me
> and to write a story about it.

No additional directions, either verbally or by example, were given. Each child chose a piece of paper and selected his/her desired writing instrument and began the task. After completing the drawing, each child selected a second piece of paper and began the second part of the task, writing the story. While the child worked, written observations were made by the researcher. For each succeeding data collecting session the above remained the same except the directions were not repeated.

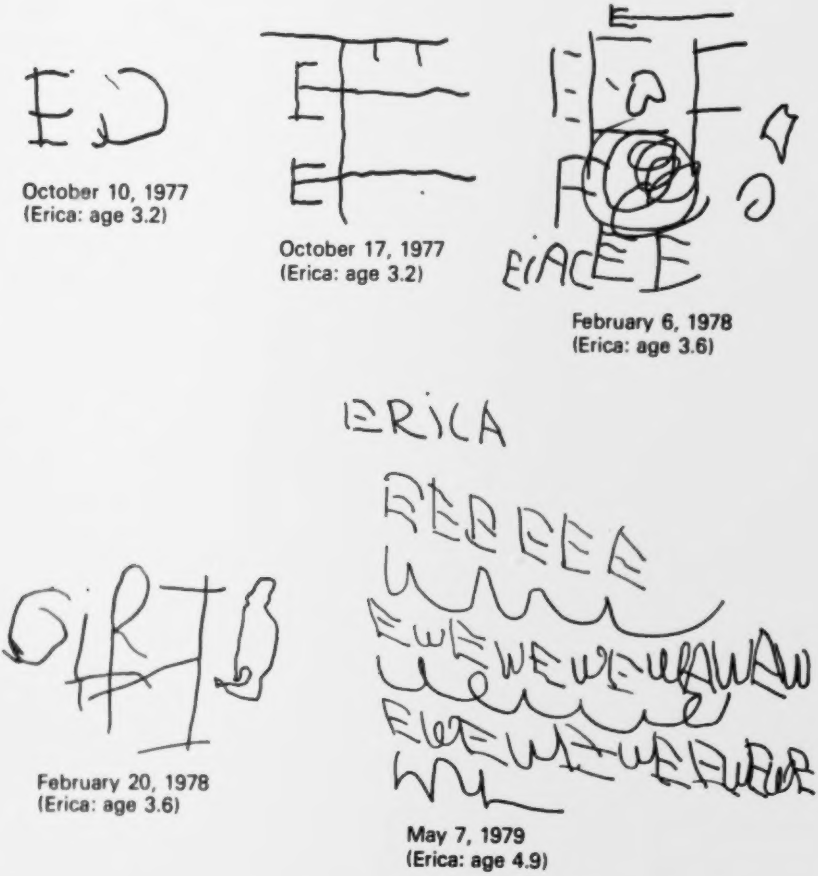
During the second year of the study the format continued with the following additions made to the task: (1) tell a story, (2) write a story and read it, and (3) write a letter, address an envelope, and read the letter and envelope. Materials remained the same, although an audio cassette player was placed on the table in order to have a record of each child's oral stories and readings. No other directions were given other than to show the children a blank envelope when asking them to write the letter. This was done to give them a context for the word letter, a message rather than a symbol.

Analysis of Data

To conceptualize the data, selected samples of Erica's writing will be used for discussion purposes. This decision has been made in order to show one child's writing over time although the data collected from the other children bear up the assumptions as previously mentioned. Erica's writing was selected because it represents the broadest age range and she had no formal instruction as did Alison and Megan.

By viewing only her product it can be noted that Erica is aware that written language has form. (See Figure 1.) Erica wrote left to right, top to bottom, and formed a number of symbols including all the letters in her name as well as F and G which closely resemble the E and C of her name. Erica is aware of alternative scripts as she initially wrote using manuscript (see October 10, 1977) then several months later began to experiment with cursive script (see May 7, 1979).

Erica's in-process behavior indicates that she can talk about written language as well as produce it. While writing on October 17, 1977, she



October 10, 1977
(Erica: age 3.2)

October 17, 1977
(Erica: age 3.2)

February 6, 1978
(Erica: age 3.6)

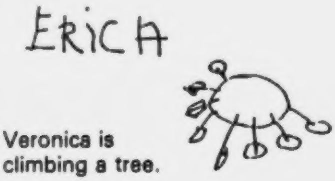
February 20, 1978
(Erica: age 3.6)

May 7, 1979
(Erica: age 4.9)

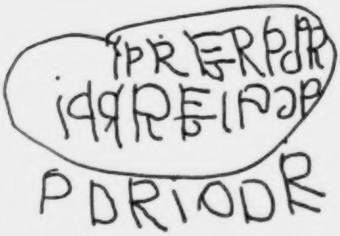
Figure 1. Writing samples demonstrating awareness that written language has form. (Reduced from the originals.)

commented that she made E's and that they were funny letters. On February 6, 1978, she remarked that she knew how to write all of her name, although she did not do so on that day. Instead she wrote F's and G's while saying, "There's an F. I know about G's." She then wrote E, i, A, C and identified each symbol. She identified the i as small i indicating her awareness of upper and lower case letters. When writing the first symbol on February 20, 1978, she began by writing C and identifying it as such. Then she said, "No, it was a G," whereupon she extended the C to form G. The symbol on the right was not identified, but she commented that it looked like an S. Erica's behavior while writing the many E's on May 7, 1979, was similar to that of October 17, 1977, when she experimented with the letter E. As Erica wrote the E's she remarked, "Oh look at all the E's. I like to make E's." In both instances she enjoyed the process of writing and shared it with this observer.

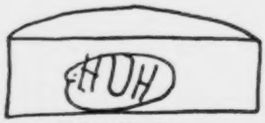
Erica has been selective in extracting from her print environment that print which is most meaningful to her. Initially her writing consists of symbols which closely resemble those found in her name. Thus from E she writes F and P, from C a G and an O, from A a V and an H. As a departure



transcript: Veronica is climbing a tree.



October 10, 1978
(Erica: age 4.2)



transcript: Hugh

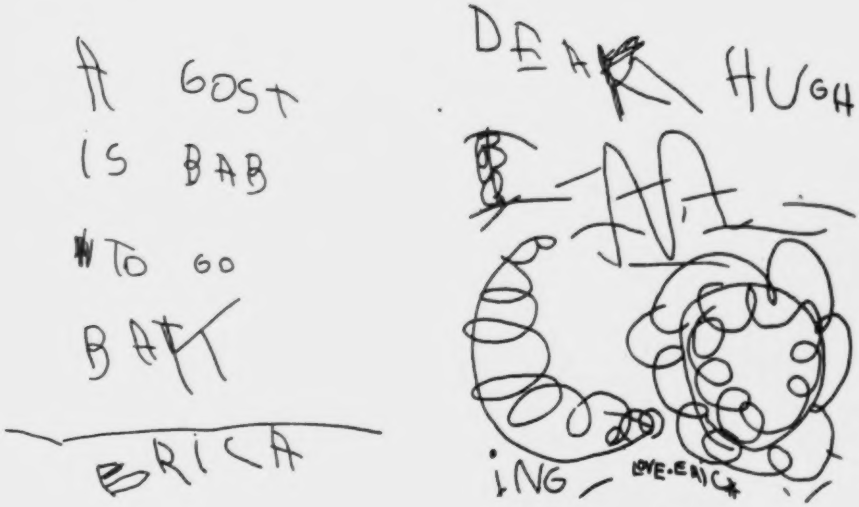
October 10, 1978
(Erica: age 4.2)

Figure 2. Writing samples demonstrating awareness that the purpose of print is to communicate. (Reduced from the originals.)

from Erica's writing, Megan, Alison, and Nicole did similarly. For example Megan wrote Z's which resemble the N of her name, Alison wrote U's which can be adapted from the O, and Nicole can add a line to the N of her name to form the M.

Erica's letter and envelope are examples of her writing which indicate she is aware that the purpose of written language is to communicate and that it has different functions. (See Figure 2.) While observing Erica, it was noted that in writing the letter she drew a picture and then wrote the first two lines of print, embellishing the E's and the lower-case d. After she had completed "experimenting" with the letters she enclosed them and wrote the last line of print, identifying each symbol while doing so. Erica then addressed the envelope, again using a string of symbols and encircling them. By viewing her product it can be recognized that the function of the letter and the function of an envelope determine the form of the print.

When Erica was asked to read her letter she pointed to the picture and said "Veronica is climbing a tree." Then Erica pointed to the symbols and said, "Well, I don't know what this says. Those are just letters." When asked to read what she had written on the envelope, Erica responded, "That says Hugh." Erica seems to be aware that it is the language user who brings meaning to the text. She has done this when she uses art to negotiate the print in order to communicate and has written an individual's name on the



May 7, 1979
(Erica: age 4.9)

transcript: A ghost is bad
to go back.

May 7, 1979
(Erica: age 4.9)

transcript: Dear Hugh
I am going
Love. Erica

Figure 3. Writing samples demonstrating awareness that written language has different functions. (Reduced from the originals.)

envelope. The string of symbols remains only a group of letters which indicates her awareness that letters in and of themselves have no inherent meaning. In so doing Erica indicates her awareness that the purpose of written language is to communicate and that it has both function and form.

The third assumption that this researcher makes is that young children are not only aware of the mechanical aspects of writing, but are also cognitively aware of different functions for writing. Each child knew that stories have a structure that is different from the structure of letters. That is, they were aware that stories have a structure (Stein, 1978) and that letters have a different kind of structure. Stories typically began with the familiar "Once upon a time" with a character/characters then introduced, followed by a situation and then a conclusion. In the majority of stories the conclusion was followed by a closure such as the end and that's all. Letters had greetings; a message, either a request or the giving of information; and a closure with a signature. When each of the children read their stories and letters they read their stories as stories and their letters as letters. They thus indicate their awareness that story, letter, and envelope writing all serve different purposes. In each task the children selected the format and text appropriate to the function. In Figure 3 it is noted that Erica's writing is different both as to function and as to form. She has written a story which functions and looks like a story and has written a letter which functions and looks like a letter.

Implications

The results of this study suggest the importance of observing children (1) in the writing process, (2) under different contextual constraints, and (3) over an extended period of time. When only the product and not both the product and the process are observed and evaluated, information about a child's knowledge of written language is limited. Just as the reader of this study has limited information about the children's print awareness since he/she does not have the advantage of having observed the process. The reader must instead rely on the product alone. Although when it is not possible to observe the process, assumptions can be made about children's print awareness when comparing writing which has been produced under different contextual constraints and when comparing writing which has been produced over a period of time. By so doing, we can recognize the natural development of young children's print awareness.

In recognizing that written language is developed naturally we can question assumptions currently undergirding writing instruction in many first-grade classrooms and provide options for instruction. Of importance is the need to question those instructional practices which assume that (1) writing is not a natural process but instead is a skill which needs to be taught; (2) writing proceeds from form to function; one has to be taught how to form letters, match up sounds and symbols, and writing simple words before expressing meaningful discourse; and (3) special materials are necessary such as lined paper, directional arrows, large sized pencils and models to trace over and under.

As an alternative to learning that writing is a mechanical skill, children need to learn that writing is a natural process. To do this they need to be in an environment which fosters using written language in meaningful ways. That is children should write and draw stories, write letters, write class helper lists, write slogans for bulletin boards, and author books but to name a few activities. Children need to see others write and to share their own writing with others. Children need to have a variety of audiences and not just be limited to the teacher as audience. Children need to have access to pencils, crayons, pens, paper and a place to write without interruption. Evaluation of children's writing needs to be made on a variety of material and seen as an on-going developmental process dependent upon function rather than upon form alone.

REFERENCES

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Framework for comprehending discourse* (Tech. Rep. No. 12). Urbana-Champaign, Ill.: University of Illinois, Laboratory for Cognitive Studies in Education, 1976.

Chomsky, C. Reading, writing, and phonology. *Harvard Educational Review*, 1970, *40*, 287-309.

Clay, M. *What did I write?* London: Heinemann Educational Books, 1975.

Harste, J. C., & Burke, C. L. Toward a socio-psycholinguistic model of reading. *Viewpoints on Teaching and Learning*, 1978, *58* (3), 9-34.

Harste, J. C., Burke, C. L., & Woodward, V. A. Children's language and world: Initial encounters with print. *Bridging the gap: Reader meets author*, in press.

Read, C. *Children's categorizations of speech sounds in English*. Urbana, Ill.: National Council of Teachers of English Committee on Research, 1975.

Stein, N. L. *How children understand stories: A developmental analysis* (Tech. Rep. No. 69). Urbana-Champaign, Ill.: Center for the Study of Reading, University of Illinois, March 1978.

PATRICIA S. KOSKINEN
LINDA B. GAMBRELL
University of Maryland

THE EFFECTS OF AN INDUCED VISUAL IMAGERY STRATEGY UPON THE READING COMPREHENSION OF BELOW AVERAGE READERS

Considerable attention has been given to the use of induced visual imagery as a strategy for enhancing memory of words and comprehension of connected discourse. Instructing children and adults to construct mental images has facilitated their learning in a wide range of situations. While researchers have carefully documented the positive effects of induced visual imagery on paired associate learning (Paivio & Yuille, 1967; Wolff & Levin, 1972; Bender & Levin, 1976), the value of using imagery as a strategy for enhancing reading comprehension has yet to be thoroughly explored. Investigations with prose have found that imagery instructions facilitate recall for adults and high school students (Steingart & Glock, 1979; Anderson & Kulhavy, 1972) as well as for average and above average readers in grades three through six (Rasco, Tennyson & Boutwell, 1975; Kulhavy & Swenson, 1975; Pressley, 1976; Levin, Guttmann & McCabe, 1977). Research results with below average readers, however, indicate that imagery instructions may only be effective with students who have average or better vocabulary skills (Levin, 1973).

While brief imagery instructions appear to facilitate listening comprehension for students as young as those in the third grade (Shimron, 1975; Guttmann, Levin & Pressley, 1977), there is conflicting information as to the type of imagery instructions or extent of training needed for students to derive benefits from the use of visual imagery in reading connected discourse. Some researchers have found brief imagery instructions were sufficient to enhance the reading comprehension of average and above average readers in the fourth through sixth grades (Kulhavy & Swenson, 1975; Rasco, Tennyson & Boutwell, 1975). Others, however, found that below average fourth grade readers with average vocabulary scores benefited from short imagery instructions while those with low vocabulary scores did not (Levin, 1973), and that brief imagery instructions did not improve the memory of fourth graders (Levin & Divine-Hawkins, 1974).

In situations where training was given in the use of visual imagery, Lesgold, McCormick & Golinkoff (1975) found a four week training program successful in helping third and fourth graders recall more information. Similar facilitation effects were found with third graders who were given a considerably shorter 20 minute visual imagery program (Pressley, 1976).

In light of the inconsistent findings reported by researchers who used only brief imagery instructions and the success of those who used an imagery training program, it appears that training may be an important variable in enabling some readers to use visual imagery. Pressley (1976) noted a trend in his data which indicated that below average readers could benefit from this training.

Since there is limited information on the use of imagery with students who are experiencing difficulty in learning to read, there is a need to further explore this area. The present study investigated the viability of using induced visual imagery as a strategy for the enhancement of reading comprehension for below average sixth grade readers. Because it appears that visual imagery may provide an organizational framework which can aid memory, a major purpose of this study was to determine the effectiveness of one type of induced visual imagery training in enabling to enable these readers to improve their reading comprehension.

A further purpose of this study was to determine if some below average readers benefit more from the use of imagery than others. Specifically, we were interested in whether visual imagery was differentially effective for students of different vocabulary levels as was found by Levin (1973). Vocabulary scores were selected as a variable to investigate based on Wiener and Cromer's work (1967) which differentiated between below average readers whose comprehension was low because they lacked decoding or vocabulary skills, and those who had these skills but had difficulty integrating text.

METHOD

Subjects

Seventy-two sixth grade students enrolled in seven suburban Maryland public schools were randomly assigned to the experimental and control conditions. All subjects met the following criteria: (1) *Iowa Test for Basic Skills* reading comprehension score between the 4th and 39th percentile, (2) *Iowa Test for Basic Skills* vocabulary score at or above the 4th percentile, and (3) *Cognitive Abilities Test* score within 1 standard deviation plus or minus the mean. Teachers also verified that subjects were correctly identified as below average readers, reading below the sixth grade level but at least on a third grade reading level.

Materials

The materials, similar in structure and format to those developed by Pressley (1976), included (a) two sentences which contained two major elements (e.g., the tiger *licked* his paws.), (b) four sentences which contained three major elements (e.g., The *dog* walked into the *box* near the *tree*.), (c) two short paragraphs approximately 20 and 40 words long, and (d) part of a short story approximately 250 words in length. The story was divided into meaningful sections and bound into a booklet. Each printed page was followed by a blank page so students would be encouraged to remember, not while reading, but immediately afterwards. The objects and activities referred to in the training materials were illustrated by pen and ink drawings and shown to the subjects in the experimental condition.

The two test stories, approximately 950 words in length, were written at a third grade readability level as determined by the Fry Readability Formula (1968). They were each divided into 14 sections, each section followed by a blank page, and bound into two booklets.

For each test story a set of 20 short answer test questions was constructed. Ten literal and ten inferential questions were asked for each test story.

Procedure

Subjects participated in small groups of three to five children. The training procedures in each group took approximately 20 minutes.

Experimental group training procedure. Subjects in the experimental conditions were given training in inducing visual imagery. The subjects were orally presented six sentences and two paragraphs, and instructed after each "to make a picture in your head to help you remember". After each sentence or paragraph the experimenter then showed the subjects a picture depicting the major elements of the prose. Subjects were told that the picture was what their mental image *might* look like and were asked if their picture was like the one presented. Subjects were told that their picture did not have to be exactly like the illustration, just so long as it contained all the same elements. Each subject then received a copy of the training story booklet. The subjects were instructed to read the first page, turn to the subsequent blank page, and "make a picture in your head to help you remember." Following the reading of the training story, the subjects were shown pictures depicting the major events of the story and then asked questions about the story.

Control group training procedure. Subjects in the control condition were given the same materials and training as the experimental group. The experimenter told the control subjects to "do whatever you can to help you remember." Control subjects did not see the pictures used in the experimental condition. Instead, subjects were instructed to recall the item, and accuracy of the recall was discussed with the subjects.

Test sessions. After training, the subjects read the first test booklet and experimental subjects were instructed to construct mental pictures to help them remember while looking at the blank pages, while control subjects were told to do whatever possible to help them remember (writing or drawing was not allowed). These instructions were repeated after three and six minutes. The amount of time each student took to read the story was recorded. Subjects were forewarned that they would have questions to answer following the reading of the story and then, after reading, given a 20 item question-booklet to assess their immediate recall. Seven days later both experimental and control subjects were given the same 20 questions to assess delayed recall of the first test story. They were then reminded to use either imagery or a general memory strategy while they read a second test story. The amount of time each student took to read the second story was also recorded. To assess the effects of training over time, the subjects were asked to answer another set of 20 questions related to the content of the second story.

RESULTS

The mean number of correct answers for immediate cued recall (Test Story 1), the delayed recall (Test Story 1) and immediate cued recall (Test Story 2) are recorded in Tables 1, 2, and 3. In all cases the mean number of correct answers was based on a total of 20 questions, 10 literal and 10 inferential. The multivariate analysis of variance and covariance procedures were employed to test the overall null hypothesis of no difference between the dependent variables of immediate and delayed recall due to treatment, sex of subject, and question level. Vocabulary level was the

TABLE 1

Means and Standard Deviations of Number of Correct Answers to Literal and Inferential Questions for Test Story 1 — Immediate Cued Recall

		Experimental		Control	
		Literal	Inferential	Literal	Inferential
Correct Answers					
Boys	M	5.050	6.050	6.000	5.500
	S D	2.164	2.235	1.862	2.066
Girls	M	3.947	4.368	4.882	4.706
	S D	2.392	1.978	1.576	1.759

TABLE 2

Means and Standard Deviations of Number of Correct Answers to Literal and Inferential Questions for Test Story 1 — Delayed Cued Recall

		Experimental		Control	
		Literal	Inferential	Literal	Inferential
Correct Answers					
Boys	M	4.600	5.450	4.562	5.313
	S D	2.583	2.373	2.308	2.600
Girls	M	3.421	4.158	4.000	4.824
	S D	1.895	2.198	1.541	1.845

TABLE 3

Means and Standard Deviations of Number of Correct Answers to Literal and Inferential Questions for Test Story 2 — Immediate Cued Recall

		Experimental		Control	
		Literal	Inferential	Literal	Inferential
Correct Answers					
Boys	M	4.900	5.650	6.000	6.125
	S D	2.573	2.477	2.066	1.204
Girls	M	4.000	5.316	5.000	5.647
	S D	1.826	2.262	2.398	1.730

covariable. The *F*-statistic for treatment was not statistically significant at the $p < .05$ level. The resulting *F*-statistics for sex, level of question, and interaction effects were not statistically signficant. An analogous analysis of variance was conducted on the reading time for the experimental and control conditions. There were no significant effects found in this analysis.

DISCUSSION

The results of this experiment are not consistent with the findings of the few other investigations which have explored the use of imagery as a means of enhancing comprehension. Imagery instructions in this study did not facilitate the performance of below average readers relative to the control group. These findings may be attributed to a number of factors. One factor may be that since control subjects were instructed to "do whatever is necessary to remember", they may have been imagining spontaneously. Levin and Divine-Hawkins (1974) mention this as a possibility as does Pressley (1976). While Pressley did get significant differences between his experimental and control conditions using similar procedures, these differences accounted for only 6% of the variance (Pressley, 1977).

The amount of imagery training these below average readers received may be another important factor. While simple imagery instructions may enhance the learning of words, considerable training may be necessary to facilitate the information processing required in reading connected discourse. Research by Shimron (1975) and Guttmann, Levin and Pressley (1977), indicate that age is important in the development of imaginal mediation of prose. In their comparison of primary children of different ages, they found that while kindergarten, first and second grades did not benefit from imagery instructions, the listening comprehension of third and fourth graders was enhanced when these children were given brief instructions to make mental images. It may be that average and above average students may benefit from brief training, but older students who are experiencing difficulty in learning to read may need extensive training.

The amount and complexity of the prose to be read may also be a factor contributing to whether visual imagery enhances memory. Some of Levin's (1973) subjects benefited from brief imagery instructions, but the material they read consisted of 12 short, high imagery sentences. The material used in the present study, however, was image evoking and similar in length and structure to the stories in children's readers (approximately 950 words). As prose selections become longer and more complex, it may be that training needs to focus not only on the interrelationships of words and sentences, but that more emphasis needs to be given to story structure.

REFERENCES

Anderson, R. C., & Kulhavy, R. W. Imagery and prose learning. *Journal of Educational Psychology*, 1972, *63*, 242-243.

Bender, B. G., & Levin, J. R. Motor activity, anticipated motor activity, and young children's associative learning. *Child Development*, 1976, *47*, 560-62.

Fry, E. A. Readability formula that saves time. *Journal of Reading Behavior*, 1968, *11*, 513-516, 575-578.

Guttmann, J., Levin, J. R., & Pressley, G. M. Pictures, partial pictures, and young children's oral prose learning. *Journal of Educational Psychology*, 1977, *69*, 473-480.

Kulhavy, R. W., & Swenson, I. Imagery instructions and the comprehension of text. *British Journal of Educational Psychology*, 1975, *45*, 47-51.

Lesgold, A. M., McCormick, C., & Golinkoff, R. M. Imagery training and children's prose learning. *Journal of Educational Psychology*, 1975, *67*, 663-667.

Levin, J. R. Inducing comprehension in poor readers: A test of a recent model. *Journal of Educational Psychology*, 1973, *65*, 19-24.

Bender, B. G., & Levin, J. R. Motor activity, anticipated motor activity, and young children's associative learning. *Child Development*, 1976, *47*, 560-62.

Fry, E. A. Readability formula that saves time. *Journal of Reading Behavior*, 1968, *11*, 513-516, 575-578.

Guttmann, J., Levin, J. R., & Pressley, G. M. Pictures, partial pictures, and young children's oral prose learning. *Journal of Educational Psychology*, 1977, *69*, 473-480.

Kulhavy, R. W., & Swenson, I. Imagery instructions and the comprehension of text. *British Journal of Educational Psychology*, 1975, *45*, 47-51.

Lesgold, A. M., McCormick, C., & Golinkoff, R. M. Imagery training and children's prose learning. *Journal of Educational Psychology*, 1975, *67*, 663-667.

Levin, J. R. Inducing comprehension in poor readers: A test of a recent model. *Journal of Educational Psychology*, 1973, *65*, 19-24.

Levin, J. R., & Divine-Hawkins, P. Visual imagery as a prose learning process. *Journal of Reading Behavior*, 1974, *6*, 23-30.

Levin, J. R., Guttmann, J., & McCabe, A. E. Distinguishing between two covert learning strategies in elementary school children. *American Educational Research Journal*, 1977, *14*, 339-342.

Paivio, A., & Yuille, J. C. Mediation instructions and word attributes in paired-associate learning. *Psychonomic Science*, 1967, *8*, 65-66.

Pressley, G. M. Mental imagery helps eight-year olds remember what they read. *Journal of Educational Psychology*, 1976, *68*, 355-359.

Pressley, M. Imagery and children's learning: Putting the picture in developmental perspective. *Review of Educational Research*, 1977, *47*, 585-622.

Rasco, R. W., Tennyson, R. D., & Boutwell, R. C. Imagery instructions and drawings in learning prose. *Journal of Educational Psychology*, 1975, *67*, 188-192.

Shimron, J. Imagery and comprehension of prose by elementary school children. (Doctoral dissertation, University of Pittsburgh, 1974.) *Dissertation Abstracts International*, 1975, *36*, 795-A. (University Microfilms No. 75-18, 254)

Steingart, S. R., & Glock, M. D. Imagery and the recall of connected discourse. *Reading Research Quarterly*, 1979, *15*, 66-83.

Wiener, M., & Cromer, W. Reading and reading difficulty: A conceptual analysis. *Harvard Educational Review*, 1967, *37*, 620-643.

Wolff, P., & Levin, J. R. The role of overt activity in children's imagery production. *Child Development*, 1972, *43*, 537-547.

JUDITH A. LANGER
East Meadow Public Schools
New York University

RELATION BETWEEN LEVELS OF PRIOR KNOWLEDGE AND THE ORGANIZATION OF RECALL

Background

Prior knowledge and its effect on the recall of text have recently become the focus of a number of reading research studies. Findings suggest that it is the reader's prior knowledge which permits anticipation of the author's intended message and leads to the comprehension of textbook material (Adams & Collins, 1979). New ideas and information are learned and retained most efficiently when particularly relevant ideas are already available within the reader's cognitive structures. The organizational structure of knowledge facilitates learning and remembering of information (Anderson, Spiro, & Anderson, 1978), and may provide a plan which helps readers retrieve information (Pichert & Anderson, 1977).

Pearson, Hansen and Gordon (1979) suggest that comprehension involves the integration of new information with preexisting schemata. If the schemata are weakly developed, comprehension requiring the integration of new and known information is difficult. Tannen (1979) found that anticipatory structures which support the comprehension and processing of stories also serve to filter comprehension and influence recall. Anderson and Freebody (1979), in their review of the role of vocabulary knowledge in reading comprehension, conclude that word knowledge is a prerequisite for reading comprehension and suggest that the development of improved methods for assessing the breadth of vocabulary knowledge is needed.

This study is based on the view that text specific concepts and vocabulary knowledge affect the processing and recall of text, and that a measure of specific text related knowledge might assist teachers in determining: 1) whether a reader possesses adequate background knowledge to comprehend and recall from textbook reading, or 2) whether additional concept instruction is necessary before textbook reading is assigned.

The purpose of this study was to elicit prior knowledge, using free association, to categorize this knowledge into broad levels, and to determine the extent of the relationship between these assigned levels of prior knowledge and the organization of recall. It is hoped the prior knowledge measure used in this study might eventually be refined for classroom teachers to easily use to predict whether a student possesses sufficient knowledge to successfully process and recall the information contained in a specific text.

METHOD

Subjects

The subjects for this exploratory investigation were 36 high school seniors from a middle class suburban school district in Long Island, New York. All were college bound students enrolled in an advanced placement course in English literature.

Procedure

The subjects were told that they would be asked to free associate with stimulus content words selected from two passages they would later be asked to read and recall. The selected passages were two of the three passages Meyer (1975) manipulated and analyzed for her work on the organization of prose and the structure of recall. The passage topics dealt with Schizophrenia and Parakeets and the with-signaling target paragraph high version of each was used.

Three content words were selected from the top half of each passage structure and used as stimuli for free association. The subjects were given each content word separately and told to write anything that came to mind when they heard the word. The free association stimulus was used to access from memory available knowledge related to the content word.

After the three words for a passage had been given and all free associations elicited, the subjects read the passage silently and then wrote all they could remember about the passage.

Analysis

The levels of prior knowledge assigned to the free association responses were categorized based on pilot study findings (see Figure 1) and were weighted from 3 to 1 with 3 representing much prior knowledge and 1 representing little prior knowledge. The scores for responses to each passage's three stimulus content words were averaged. Actual prior knowledge score averages ranged from 2.67 to 1.00.

MUCH — (3)	**SOME — (2)**	**LITTLE — (1)**
superordinate concepts, definitions, analogies, and linking	examples, attributes, and defining characteristics	associations, morphemes, sound alikes, and first-hand experiences

FIGURE 1 Levels of Prior Knowledge

The recall protocols were scored according to Meyer's text analysis categories which were based on Fillmore's (1968) case grammar and Grimes' (1972) semantic staging of propositions (see Tables 1 and 2). One point was given for inclusion in the recall of content words, lexical predicates, role relations and rhetorical predicates. Meyer's hierarchically structured passage content was divided into thirds and items included in each of the three height levels were computed separately. The content words, lexical predicates and roles were then scored one point while the rhetorical predicates, which represented the superordinate structures, were weighted two points. Finally, a total summed recall score was computed with each of the first four variables assigned one point.

Two judges scored the prior knowledge and recall items separately. Interrater agreement revealed a relatively high percentage of agreement on the recall scores (.94) and somewhat lower on the levels of prior knowledge scores (.82). In cases of disagreement, an average of the two scores was computed for the prior knowledge measure, and the decisions were negotiated for the recall analysis.

A Pearson Product Moment Correlation was carried out to determine the extent of the relationship between the levels of prior knowledge and the aspects of recall derived from the recall analysis for each passage.

RESULTS

Analysis of the Schizophrenia recall (see Table 1) indicates that variables 2 through 10 are all highly correlated with levels of prior knowledge at $p<.05$.

The intercorrelations among variables tend to be moderately high (.351 to .946) among those which are significant. However, total responses in the middle and bottom heights of the passage structure (V7 and V8) have the lowest correlation with prior knowledge and the lowest intercorrelations with other aspects of recall as well. Therefore, although responses in the top third of the structure tend to correlate with level of prior knowledge and the various measures of organization of recall, responses falling in the lower two thirds of the passage structure appear least predictable in terms of measures of recall.

The recall measures in the Parakeet passage (see Table 2), indicate a high correlation with levels of prior knowledge at $p<.05$. The intercorrelations among variables tend to be moderately high (.342 to .882) among those which are significant. However, V18, the variable representing recall in the lowest third of the passage structure, has the lowest correlation with prior knowledge and the lowest intercorrelations with the other measures of recall.

A review of both tables indicates that although total recall correlated most highly (.741) with prior knowledge for the Schizophrenia passage, the correlation was not as high (.440) for the Parakeet passage. However, the correlation between levels of prior knowledge and content word recall remained consistently high (.72, $p<.001$) across both passages. It is also interesting to note the high level of consistency between content words and rhetorical predicates across both passages.

The variation in total recall across both passages might suggest a greater passage dependence than the total data might at first indicate. However, the protocols for the two highest and lowest scorers across both passages were examined and in these cases prior knowledge and total scores were consistent across passages.

DISCUSSION

These findings suggest that the levels of prior knowledge utilized in this study are highly correlated with the reader's organization of recall, and that the prior knowledge categories might be used as predictors of recall success prior to the assignment of passage reading in the classroom. The low correlations can be explained by examining Meyer's (1975) conclusion that information high in the content structure is recalled and retained better than information at the lower levels. Also, the pattern of relations at the top of the content structure influence recall while those low in the structure have no influence on recall. Since the low correlations in this study

TABLE I

Product Moment Correlation Coefficients Between Level of Prior Knowledge and Aspects of the Organization of Recall

Schizophrenia Passage, N = 36

	V1 Level of Prior Knowledge	V2 Content Words	V3 Lexical Predicates	V4 Roles	V5 Rhetorical Predicates	V6 Responses Top Third	V7 Responses Middle Third	V8 Responses Bottom Third	V9 Total Responses Rhetorical #2	V10 Total Responses
V1	1.00	.723*	.630*	.578*	.613*	.543*	.387*	.426*	.718*	.741*
V2			.501*	.651*	.720*	.598*	.564*	.423*	.865*	.899*
V3				.696*	.539*	.592*	.239*	.500*	.649*	.730*
V4					.656*	.471*	.502*	.560*	.767*	.841*
V5						.578*	.609*	.351	.867*	.882*
V6							−.041*	.241	.648*	.659*
V7								−.080	.545*	.607*
V8									.590*	.498*
V9										.946*
V10										

*$p \leq .05$

TABLE 2

Product Moment Correlation Coefficients Between Level of Prior Knowledge and Aspects of the Organization of Recall

	V11 Level of Prior Knowledge	V12 Content Words	V13 Lexical Predicates	V14 Roles	V15 Rhetorical Predicates	V16 Responses Top Third	V17 Responses Middle Third	V18 Responses Bottom Third	V19 Total Responses Rhetorical #2	V20 Total Responses
V11	1.00	.721*	.549*	.593*	.620*	.536*	.550*	.354*	.629*	.440*
V12			.776*	.878*	.728*	.723*	.733*	.545*	.872*	.672*
V13				.747*	.554*	.570*	.652*	.369*	.648*	.519*
V14					.599*	.604*	.682*	.596*	.798*	.596*
V15						.606*	.748*	.224	.763*	.590*
V16							.342*	.156	.644*	.435*
V17								.216	.778*	.653*
V18									.612*	.621*
V19										.882*
V20										

*$p \leq .05$

correspond with the variables representing the middle and bottom third of the content structure, it can be concluded that the information located at the lower levels of the content structure are not related either to prior knowledge or the other recall variables.

Since the positive relationship between the levels of prior knowledge and recall has been demonstrated, future research might more fully examine whether difference in prior knowledge level makes a difference in what is retained. Due to the number of variables included in this study, replication with a larger number of subjects is desirable. Further research might more specifically investigate the influence of prior knowledge on total responses, content words, and rhetorical predicates in the recall protocols. An analysis of items included in the recall which had not appeared in the text might provide additional insights into the relationship between prior knowledge and recall.

REFERENCES

Adams, M. J., & Collins, A. A schema theoretic view of reading. In R. O. Freedle (Ed.), *New directions in discourse processing* (Vol. II). Norwood, New Jersey: Ablex, 1979.

Anderson, R. C., & Freebody, P. *Vocabulary knowledge*. (Tech. Rep. No. 136). Champaign, Ill.: Center for the Study of Reading, University of Illinois, August 1979.

Anderson, R. C., Spiro, R. J. & Anderson, M. C. Schemata as scaffolding for the representation in connected discourse. *American Educational Research Journal*, 1978, *15*, 433-450.

Fillmore, C. J. The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart & Winston, 1968.

Grimes, J. E. *The thread of discourse*. The Hague: Mouton, 1975. (ERIC Document Reproduction Service No. ED 071 485)

Meyer, B. J. F. *The organization of prose and its effects on memory*. Amsterdam: North Holland, 1975.

Pearson, P. D., Hansen, J. & Gordon, C. The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior*, 1979, *II*, 201-209.

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, *69*, 309-315.

Tannen, D. What's in a frame? Surface evidence for underlying expectations. In R. O. Freedle (Ed.), *New directions in discourse processing* (Vol. II). Norwood, New Jersey: Ablex, 1979.

DR. JOHN M. MIZE
Alamance County Schools
Graham, N. C. 27253

THE RELATIONSHIP BETWEEN READING PERFORMANCE AND AWARENESS OF COMPREHENSION OF TEXTBOOK PASSAGES

Recently, researchers have begun to examine the student's awareness of the processes employed to read and comprehend. Much of the work that has been done suggests Mattingly's concept of "linguistic awareness"—the awareness of a child of the rules of the language he uses. Mattingly (1972) suggested that in order for a child to be an effective language user he must be familiar with these rules. Downing (1979) extended this notion in his theory of reading acquisition which centers around the notion of "cognitive clarity"—a condition necessary for efficient language use and in particular for efficient reading. According to Downing, cognitive clarity is achieved as the child learns that language and the manipulation of language can be an object of study rather than an automatic process over which one exerts no control.

As the child learns about language, monitoring processes direct learning and allow the learner to make the decision that learning has been accomplished (Mattingly, 1972; Flavell, 1977; Gagne, 1977; Downing, 1979). Consequently, if comprehension is learning the meaning of the passage, then the reader must use these monitoring processes to make decisions concerning comprehension or non-comprehension of the passage which is being read. A decision that one has comprehended would send a reader forward to new material; a decision that one has not comprehended might cause a reader to employ appropriate strategies to remedy his non-comprehension.

Several studies (Frase & Silbiger, 1970; Frase, 1975; Hagen & Stanovich, 1977; Kobasigawa, 1977) suggest that good readers monitor their own reading in such a way that they can make decisions regarding the most appropriate strategy to use. If this inference is correct, then it would seem that good readers should have a greater awareness of whether they have comprehended what they have read than poor readers would have. Consequently, one might speculate that a good reader continually asks himself whether he is understanding what he is reading. This might well be the first question he asks himself, because his answer determines if and how strategies are to be applied.

The present study is designed to investigate if reading ability is related to a reader's awareness of whether he is comprehending a text passage as he reads. Since one cannot directly measure awareness of comprehension, the study was designed to investigate if reading ability as measured by a standardized reading test is related to a reader's ability to make judgements concerning the correctness or incorrectness of his answers to comprehension questions. It is assumed that awareness of comprehension is reflected in a reader's ability to make judgements concerning the correctness or incorrectness of his answer. This topic is chosen for investigation because, as stated previously, it seems that an accurate judgement of comprehension or non-comprehension is what determines if and how strategies are applied.

METHOD

Subjects and Procedures

The subjects were 144 seventh graders who attended a middle school in a central Georgia town. Each subject was asked to predict how many of the fourteen questions he would answer correctly before reading each of two passages (the confidence index). After reading the passages, each subject was again asked to predict the number of questions he would answer correctly (the index of general judgement). For every item of the test, each subject was asked to judge whether he had answered the item correctly or incorrectly (the index of specific judgement). While answering the questions, half the students were allowed to refer to the passage; half were not (the passage condition). Following the test, students were randomly chosen to be interviewed.

Materials

The Iowa Tests of Basic Skills: Test R, Reading Comprehension (Linquist & Hieronymous, 1964) was used as the measure of the reading comprehension ability of each subject. The passages which were read by the subjects and over which testing occurred were selected from a set of cause-effect passages randomly selected from six basal readers, six science texts, and six social studies texts, all on the Georgia State-Approved Textbook list. The test passages were each approximately 315 words in length, with an index of readability of high sixth grade as determined by a formula contrived by Dale and Chall (1948).

The test of comprehension for each passage consisted of fourteen multiple-choice items. The first ten questions assessed literal recall of information relating to the central cause-effect relationship of the passage. The last four questions required more than literal recall and assessed comprehension of the passage's main idea. The coefficient of internal consistency (Cronbach, 1951) for each passage was .70.

Design

Two measures of ability to judge degree of comprehension were the index of specific judgement and the index of general judgement. The former was related to reading ability, the passage condition, and the number of correct responses. The latter was related to reading ability, the passage condition, the number of correct responses, and the confidence index. The independent variables were subjected to a stepwise regression analysis to determine how each set predicted the variance in the respective dependent variables. Hierarchical regression analyses were conducted and zero order correlation coefficients were computed as a *post hoc* examination of the data.

RESULTS

The statistical analysis of the data yielded the following results.

1. The number of correct responses (M=17.23, SD=5.37) best predicted ($p<.05$) the index of specific judgement (M=18.52, SD=4.62). Reading

ability and the passage condition did not account for a significant amount of variance in the index of specific judgement.

When placed first in a prediction equation, reading ability accounted for 12 percent of the variance ($p < .05$) in the criterion measure. When placed second, the passage condition accounted for 25 percent of the variance ($p < .05$). When placed third, the number correct accounted for 35 percent of the variance ($p < .05$). While all three variables accounted for a significant proportion of variance in the dependent variable when entered in this order, the number correct still accounted for the most variance and reading ability the least.

There was a positive relationship of .44 between reading ability and the number correct and a positive relationship of .36 between reading ability and the index of specific judgement. A very strong relationship of .86 existed between the number correct and the index of specific judgement.

2. The confidence index ($M=22.83$, $SD=3.60$) best predicted ($p < .05$) the index of general judgement ($M=23.22$, $SD=3.83$). Reading ability also contributed to predicting ($p < .05$) the index of general judgement but accounted for only 2 percent of the variance. In comparison, the confidence index accounted for 56 percent of the variance.

When placed first in a prediction equation, reading ability accounted for 16 percent of the variance ($p < .05$) in the criterion measure. When placed second, the number correct accounted for 6 percent of the variance ($p < .05$). When placed third, the confidence index accounted for 37 percent of the variance ($p < .05$). When placed last, the passage condition accounted for less than 1 percent of the variance (ns). While three out of the four variables accounted for a significant proportion of variance in the dependent variable when entered in this order, the confidence index still accounted for the most variance and reading ability for a relatively small amount.

There was a positive relationship ($r=.35$) between reading ability and the confidence index and a positive relationship ($r=.40$) between reading ability and the index of general judgement. There was also a positive relationship ($r=.39$) between the number correct and the index of general judgement. There was a very strong relationship ($r=.75$) between the confidence index and the index of general judgement.

3. Subjects did not always know whether they were comprehending as they read. Self reports suggested frequent inability to judge accurately whether comprehension was occurring.

CONCLUSIONS

The findings of this study suggest that lack of awarenes of the process of comprehension is not confined to poor readers alone. Good and poor readers alike are not always aware of whether they are comprehending as they read. It appears that an individual's immediate performance is a better predictor of awareness than overall reading ability. If this is so, then awareness of comprehension may be more closely related to the content of the material read. In other words, a student's awareness of whether he is comprehending is influenced by a student's familiarity with the ideas expressed in the passage he is reading.

In addition, it appears that a student's confidence in himself as a reader directly affects to what degree a student thinks he has comprehended a passage. Very often, a student's general impression of his degree of comprehension does not correspond well to his actual comprehension as determined by test performance. This finding again suggests that readers at all levels of ability are not aware of whether they have comprehended adequately.

REFERENCES

Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, *16*, 297-334.

Dale, E., & Chall, J. S. A formula for predicting readability: Instructions. *Educational Research Bulletin*, 1948, *27*, 37-54.

Downing, J. *Reading and reasoning*. Edinburgh: Chambers, 1979.

Flavell, J. H. *Cognitive development*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

Frase, L. T. Prose processing. In G. H. Bower (Ed.), *The psychology of learning and motivation*. San Francisco: Academic Press, 1975.

Frase, L. T., & Silbiger, F. Some adaptive consequences of searching for information in a text. *American Educational Research Journal*, 1970, *7*, 553-560.

Gagne, R. M. *The conditions of learning*. New York: Holt, Rinehart and Winston, 1977.

Hagen, J. W., & Stanovich, K. G. Memory: Strategies of acquisition. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Kobasigawa, A. Retrieval strategies in the development of memory. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Linquist, E. F., & Hieronymous, A. N. *Iowa Tests of Basic Skills*. New York: Houghton Mifflin Co., 1964.

Mattingly, I. G. Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, Mass.: MIT Press, 1972.

REBECCA J. PALMER
WAYNE H. SLATER
MICHAEL F. GRAVES
University of Minnesota

THE EFFECT OF PASSAGE DIFFICULTY ON GOOD AND POOR READERS' USE OF AUTHOR'S SCHEMA IN WRITTEN RECALL PROTOCOLS

Recent theories of discourse structure have generated a great deal of research on the text organization used by writers and perceived by readers (Drum, 1978; Meyer, 1975, 1977; Meyer, Brandt, & Bluth, 1978; Rumelhart, 1975; Rumelhart & Ortony, 1977). In one of these investigations, Meyer, Brandt, and Bluth (1978) examined the extent to which good and poor ninth grade readers used the author's textual schema in organizing their written recall protocols. In this study, the students were asked to read two well-organized passages of expository text with clearly identifiable top-level rhetorical predicates and to write recall protocols. Analyses of the protocols indicated that most high ability ninth graders used the same schema for organizing their recall protocols as the authors of the passages while most low ability students did not, and that students who employed the strategy of using the author's textual schema recalled more information from the passages than those who did not.

Although these results are certainly reasonable, a question regarding the difficulty of the passages used in the Meyer et al. (1978) study arises. Meyer et al. claim that their treatment passages ("Loss of Body Water" and "Supertankers") were at reading levels appropriate for ninth graders. However, "Body Water" is rated as being at the college level by both the Dale-Chall (1948) and Fry (1972) readability formulas, and "Supertankers" is rated as being at the eleventh grade level by the Dale-Chall formula and the tenth grade level by the Fry formula. Thus, it is possible that the poorer readers in the Meyer et al. study failed to follow the author's schema in writing their recall protocols simply because the passages were too difficult for them to read.

The purposes of the present study were to replicate the Meyer et al. study, again examining the extent to which good and poor readers use the author's schema in organizing written recall protocols, to investigate the effect of passage difficulty on good and poor readers' use of the author's schema in organizing written recall protocols, to further examine the relationship between students' use of the author's schema and the amount of information they recall, and to examine the effect of text difficulty on the amount of information students recall.

METHOD

Subjects

Twenty-eight ninth graders from a heterogeneous English class in a middle-class, largely white suburban Minneapolis junior high school participated in the study. Students were divided into groups of high (n=10), average (n=10), and low (n=8) ability readers on the basis of their performance on the comprehension subtest of the *Gates-MacGinitie*

Reading Tests, Survey F (1969). Grade equivalents for the low ability group ranged from 4.5 to 9.4; those for the middle ability group ranged from 9.6 to 11.0; and those for the high ability group ranged from 11.0 to 12.9.

Materials

Materials included written directions for subjects, the two passages from the Meyer et al. study, and simplified versions of the passages. As already noted, the original version of "Body Water" was written at the college level. It contained 142 words and 103 propositions with six hierarchical levels. The original version of "Supertankers" was written at the tenth or eleventh grade level. It contained 241 words and 133 propositions with nine hierarchical levels. The re-written versions of the two passages had the same hierarchical structure as the original versions but contained simplified vocabulary and syntax. The easier version of "Body Water" was 105 words long, contained 93 propositions, and was written at the ninth to tenth grade level according to the Dale-Chall formula and the eighth grade level according to the Fry formula. The easier version of "Supertankers" was 200 words long, contained 121 propositions, and was written at the seventh grade level according to the Dale-Chall formula and the fifth grade level according to the Fry formula.

Design

The study employed a 3 x 2 x 2 design with the variables being ability (high, middle, low), passage difficulty (easy, difficult), and passage ("Body Water," "Supertankers"). The dependent measures were (1) an index of whether or not each student's recall followed the author's textual schema and (2) the percentage of propositions each student recalled.

Procedure

Four sets of two passages, counterbalanced for order of passage and passage difficulty, were constructed, and subjects were randomly assigned to one of these sets. Each subject thus read both passages, one at each level of difficulty. After reading one passage, subjects wrote down all they could recall, using complete sentences. When subjects finished writing the first protocol, they read the second passage and wrote a second protocol.

The protocols were scored using the Meyer (1975) and the Meyer et al. (1978) scoring systems. These scoring systems yielded one score indicating whether or not subjects' protocols followed the author's schema and another score indicating the percentage of propositions from the passages that were included in the subjects' recall protocols.

RESULTS

To determine the extent to which good and poor readers used the author's schema in organizing their recall protocols, a 2 x 3 contingency table showing the number of times that high, middle and low ability students did and did not follow the author's schema was constructed. This indicated that high ability students used the author's schema in 16 of their recall protocols, middle ability students used it in 13 protocols, and low ability students used it in two protocols. It further indicated that high ability students failed to use the author's schema in four of their protocols, middle ability students failed to use it in seven protocols, and low ability students failed to use it in 14 protocols. A chi square analysis revealed a significant difference ($x^2(2)=17.56$, $p<.001$) in the number of times high, middle, and low ability readers used and didn't use the author's schema. Good readers did follow the author's schema more frequently than poor readers.

To determine whether or not students followed the author's schema more frequently with the easier versions of the passages than with the more difficult versions, a 2 x 3 contingency table showing the number of times high, middle, and low ability students followed the author's schema with the easy and difficult versions of the passages was constructed. This indicated that with the difficult versions of the passages, high ability students used the author's schema in seven of their recall protocols, middle ability students used it in seven protocols, and low ability students used it in none. It further indicated that with the easy versions of the passages, high ability students used the author's schema in nine of their recall protocols, middle ability students used it in six protocols, and low ability students used it in two protocols. A chi square analysis revealed that there was no significant difference ($x^2(2)=2.05$, $p<.25$) in the number of times students at any ability level used the author's schema with the difficult versions of the passages and the number of times they used it with the easy versions of the passages.

To examine the relationship between students' use of the author's schema and the amount of information they recalled, the percentage of propositions recalled by students using the author's schema was compared to the percentage of those recalled by students not using the author's schema. This comparison showed that students who used the author's schema recalled 50% of the propositions while those that didn't recalled only 27% of them. A two-tailed *t*-test showed that this difference was significant ($t(54)=5.39$, $p<.001$).

Finally, to examine the effect of text difficulty on the amount of information students recalled, a 2 (difficulty) x 3 (ability) analysis of variance was run. Results indicated a significant effect of ability ($F(2,50)=8.46$, $p<.001$) and the Newman-Keuls test indicated that high ability students recalled significantly ($p<.01$) more propositions than low ability students (51% vs. 26%) but that the differences between the percent recalled by middle ability students (40%) and that of the other two ability groups were not significant. However, text simplification did not result in students recalling significantly more propositions ($F(1,50)=.43$, $p<.5$). Students who read the simplified versions recalled an average of 41% of the propositions, while those reading the difficult versions recalled 38% of the propositions.

SUMMARY

The results of this study strongly support the original finding of Meyer et al. (1978) that good readers use the author's schema much more often than do poor readers. While the good readers exhibited a close adherence to the author's top level structure, poor readers' recall protocols were characterized by a lack of awareness of the top-level structure and in most cases were simply unorganized lists of ideas.

Moreover, the data also strongly support the finding of Meyer et al. that students who use the author's schema recall more information than those who do not. In fact, students who used the author's schema recalled nearly twice as much as those who did not.

Finally, contrary to expectations, simplifying passages did not aid either good or poor readers. In this study, simplification of syntax resulted in a difficult sentence such as, "Oil spills also kill microscopic plant life which provides food for sea life and produces 70 percent of the world's oxygen supply," being rewritten as two sentences, "Oil spills also kill plant life. This plant life provides food for other sea life and produces 70 percent of the earth's oxygen." Simplification of difficult vocabulary resulted in such low frequency words as *convulsions* and *maneuver* being replaced by such high frequency words as *fits* and *control*, and deletion of difficult vocabulary resulted in such words as *microscopic* and *recognition* being omitted. However, this text simplification did not have a significant effect on either students' use of the author's schema or the amount of information they recalled. This finding is consonant with current theories that text difficulty results not only from syntactic complexity and vocabulary difficulty but also from students' interest and experience, conceptual complexity, cohesiveness, and levels of staging.

In summary, these results indicate that good readers use the author's schema while poor readers don't; students who use the author's schema recall much more than those who don't; and simplifying the vocabulary and syntax of a passage does not necessarily change those outcomes. Of course, these results should be interpreted with caution since the study employed a small number of subjects and only two texts, and future research should investigate the effects of similar text manipulations with more subjects and various texts.

REFERENCES

Ball, J. A. (Ed.) Supertankers. *Read: The Magazine for Reading and English*, 1975, *24* (12), 2-4.

Dale, E., & Chall, J. S. *A formula for predicting readability*. Columbus: Ohio State University, 1948.

Drum, P. A. Prose recall responses and categories for scoring. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice*. Clemson, S. C.: National Reading Conference, 1978.

Fry, E. Selecting books. *Reading instruction for classroom and clinic*. New York: McGraw Hill, 1972.

Gates-MacGinitie Reading Tests: Survey F. New York: Teachers College Press, 1969.

Meyer, B. J. F. *The organization of prose and its effects on memory*. Amsterdam: North-Holland Publishing Company, 1975.

Meyer, B. J. F. Organization in prose and memory: Research with application to reading comprehension. In P. D. Pearson (Ed.), *Reading: Theory, research, and practice*. Clemson, S. C.: National Reading Conference, 1977.

Meyer, B. J. F. What is remembered in prose: A function of passage structure. In R. O. Freedle (Ed.), *Discourse production and comprehension*. Norwood, N. J.: Ablex Publishing Company, 1977.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. *Use of author's textual schema: Key for ninth graders' comprehension*. Paper presented at the American Educational Research Association Convention in Toronto, Canada, March 1978.

Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding*. New York: Academic Press, 1975.

Rumelhart, D., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

JAMES W. PICHERT
Vanderbilt University

SENSITIVITY TO IMPORTANCE AS A PREDICTOR OF READING COMPREHENSION

The purpose of this paper was to assess whether relationships exist between children's ability to rate and recall important prose elements and a standardized measure of reading comprehension. Distinguishing important and unimportant story elements, what we will call sensitivity to importance, seems to be a facet of one's ability to use context. Items in a text become more or less important depending on the context created by the text and contexts created by the reader's purposes, interests, and background knowledge (Anderson & Pichert, 1978; Pichert & Anderson, 1977). In other words, high level organizational processes influence context, and context helps define relative prose importance. If higher order processes are part of reading comprehension, children who are sensitive to relative prose importance should manifest higher levels of reading comprehension than their not-so-sensitive peers.

The following studies assess the relationship between sensitivity to importance and a standardized measure of reading comprehension. These studies were designed to address several questions, and other facets of the studies have been reported elsewhere (Pichert, 1979). Therefore, some details have been omitted. Only those results pertinent to the relationship between sensitivity to importance and reading comprehension are reported and discussed in this paper.

EXPERIMENT 1a

This study assessed the capacity of good and poor readers (in grades three, five, and seven) to identify story information that adults regard as important to certain perspectives.

METHOD

Subjects

Forty-five third, 45 fifth, and 51 seventh grade students from a rural Illinois school district served as judges. Also participating were 46 graduate students from an educational psychology course at the University of Illinois.

Materials

A story that could be viewed in terms of two or more high level schemata was constructed. This story, hereafter called the Skipping School passage, was written to contain approximately equal numbers of features of interest to a burglar and a safety expert. For instance, a burglar would be interested in a jewelry box but uninterested in a swing that was nearly ready to fall. Presumably the reverse would be true of a safety expert.

Procedure

The procedures were straightforward: students and adults read a specially constructed story from the safety expert or burglar perspective or no directed perspective. After reading they rated the relative importance of each idea unit on a three point scale. A median split on scores on the Stanford Achievement Test defined "high" and "low" groups on the reading comprehension subtest.

RESULTS

Changes in the ability to identify important elements

The rank orders of adult ratings were used to identify quartiles of idea units from least to most important. The mean importance rating for each child for each level of importance according to the operative perspective was computed and entered into a 3(Age) X 3(Perspective) X 2(High vs. Low Standard Reading Comprehension Score) X 4 (Idea Unit Importance Level) mixed analysis of variance. There was a main effect for age. The Perspective X Idea Unit Importance interaction was also significant. Most pertinent to this paper, the interaction between reading comprehension level and idea unit importance was significant ($F(3,330)=3.9, p<.01$). Ratings increased as a function of idea unit importance for both ability groups, but the range of ratings across levels was greater for subjects of high ability, replicating Smiley, Oakley, Worthen, Campione, and Brown (1977). Apparently, sensitivity to idea unit importance develops with age. Factors such as reading ability and whether the subject is directed to take a perspective are also determinants of this skill, however.

Congruence scores

Next we computed a "congruence" score for each subject, an index of sensitivity to idea unit importance. Each subject's ratings for the 58 idea units was correlated with the mean adult ratings from the operative perspective. Subjects sensitive to importance (defined by the adult standard) should receive scores approaching +1.00 while those who were either insensitive to or unable to understand the task would receive scores approaching +0. An analysis of variance showed that subjects with high reading comprehension scores had higher congruence scores than low ability subjects, .46 and .39 respectively ($F(1,110)=10.9, p<.01$).

EXPERIMENT 1b

Experiment 1b was conducted in order to examine the relationships between the third graders' ratings of importance and their story recall. In addition to replicating the results of Experiment 1a our aim was to see whether additional measures of higher order processing would reveal substantial relationships with reading comprehension. Therefore, eight weeks after Experiment 1a had been conducted, investigators returned to have the same third grade students listen to the same story, recall it, and rate it for a second time.

RESULTS

Children's recall was scored, using lenient gist criteria, for the presence or absence of each idea unit in the story. For each subject, point-biserial correlations were computed between recall of the 58 idea units and various rating data. Collectively these will be referred to as *concordance* scores. In spite of the fact that all of the scores were very low it is worth mentioning that virtually all of them were positive.

A stepwise multiple regression analysis was performed to identify significant predictors of the reading comprehension score. Entered for each subject as predictors were decoding speed and decoding accuracy, congruence measures, concordance measures, total recall, and proportion of important ideas recalled. Significant predictors in order of entry (and Multiple *R* value) were decoding accuracy (.44), congruence of ratings with peers (.50), concordance of recall with the subject's own ratings (.54), the proportion of important units recalled (.61), and, finally, decoding speed (.65).

EXPERIMENT 2

One purpose of Experiment 2 was to determine whether we could obtain an improved understanding of the relationship between higher order knowledge, decoding skills, and reading comprehension. The results of Experiments 1a and 1b reveal a modest correlation between sensitivity to importance and reading comprehension. Such modest relationships need to be replicated. In addition, we wondered whether the ability to recall additional story elements would be related to reading comprehension scores.

New groups of third, fifth, and seventh grade subjects (*N*=148) were directed to take either a safety expert or burglar perspective or no directed perspective and read the Skipping School passage. Everyone was then given a test of word knowledge and decoding speed. After recalling the story, subjects were asked to list story elements important to their old perspective and then list those important to a new perspective.

RESULTS

Ideas remembered after initial recall

Older and brighter children remembered a greater number of additional units.

Reading comprehension as a function of sensitivity to importance

A stepwise multiple regression analysis was performed to identify significant predictors of the standardized reading comprehension score. Significant predictors in order of entry (and Multiple *R* value) were speed of decoding (.56), total recall (.64), age (.67), decoding accuracy (.70), and, finally, the correlation between recall and adult importance ratings (.71). Together these accounted for approximately 50 percent of the variance in the reading comprehension scores.

An analysis similar to the one above revealed that the correlation between reading comprehension and what was recalled barely changed when additional recall was added to initial recall. This suggests that the ability to recall additional information may bear little relationship with school performance.

DISCUSSION

While the evidence is not very strong, it weighs in favor of the position that there is more to reading comprehension than simple decoding. Good comprehenders assigned levels of importance more like adults than poor comprehenders, replicating Smiley, et al. (1977). Sensitivity to importance on the rating task ("congruence scores") was a modest correlate of a comprehension measure in Experiments 1a and 1b. Sensitivity to importance on the free recall task ("concordance scores") was also a statistically significant predictor of comprehension. Most important, these relationships held even when one or more measures of decoding skill had been partialled out.

The relationship between decoding skills and comprehension may have been truncated by our use of a brief (20 words) decoding task. Older and brighter students performed near ceiling on the decoding task. Decoding speed correlated .56 with reading comprehension, and with decoding accuracy the multiple *R* was approximately .60. Other investigators (e.g., Perfetti & Lesgold, 1976) have obtained results which suggest that these measures account for more variance. The fact remains, however, that in both experiments measures of sensitivity to importance made small, but statistically significant, contributions toward predicting reading comprehension.

Our evidence does not rule out an interpretation that sensitivity to structure may be a function of a subject's ability to rapidly access and use previously stored information. Proponents of this view would argue that since verbal ability tests indirectly distinguish rapid and slow verbal coders, and that since rapid coders are better manipulators of processed information, high verbals should recall more additional information than low. That is what we found. This finding is also consistent with Meyer's (1977) notion that good readers could distinguish more importance levels than poor readers, not because of encoding problems, but because of problems with organization and retrieval.

The results of these experiments must be replicated with materials more like those used in classrooms. Moreover, our assessment of sensitivity to importance was based on a one-item test—ratings or recall of the Skipping School passage. A larger number and variety of texts should be used.

ACKNOWLEDGEMENT

The research reported herein was supported in part by the National Institute of Education under Contract No. US-NIE-C-400-76-0116. This paper is based on a larger work submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Education in the Graduate College of the University of Illinois at Urbana-Champaign, 1978. The author gratefully acknowledges contributions from Richard C. Anderson, Ernest T. Goetz, and Larry L. Shirey. Thanks also go to the administrators, teachers, and students of the Danville and Monticello, Illinois, school districts.

REFERENCES

Anderson, R. C., & Pichert, J. W. Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 1-12.

Meyer, B. J. F. The structure of prose: Effects on learning and memory and implications for educational practice. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, New Jersey: Erlbaum, 1977.

Perfetti, C. A., & Lesgold, A. M. *Discourse comprehension and sources of individual differences*. Paper presented at the Twelfth Annual Carnegie Symposium on Cognition, Carnegie-Mellon University, Pittsburgh, Pennsylvania, May 1976.

Pichert, J. W. *Sensitivity to what is important in prose*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1979.

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, *69*, 309-315.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written vs. oral presentation. *Journal of Educational Psychology*, 1977, *69*, 381-387.

EARL F. RANKIN
University of Kentucky

SUSAN THOMAS
Fayette County Public Schools

CONTEXTUAL CONSTRAINTS AND THE CONSTRUCT VALIDITY OF THE CLOZE PROCEDURE

Since the early 1950s when Wilson Taylor (1953) introduced the cloze procedure, numerous studies have reported convincing evidence concerning the criterion-related validity of this procedure as a method of constructing reading comprehension tests. Correlations between cloze test results and measurements of comprehension of specific articles, measures of specific high level reading skills, and measures of general reading ability by standardized tests have yielded high correlations in the .70 - .90 range (Rankin, 1964). However, the cloze procedure has not been without its critics. The main thrust of critical studies questioning the validity of this procedure concerns the capability of such tests to measure high level comprehension processes because, it is alleged, the cloze procedure is only sensitive to local constraints between words a short distance from the deletions in a cloze passage and, therefore, cannot be considered as direct measures of a construct entitled "reading comprehension." Aborn, Rubenstein, and Sterling (1959) found that increasing context beyond ten words does not increase predictability. MacGinitie (1960, 1961) in an often quoted study found that increasing context beyond five words had no influence upon word predictablity in cloze passages. Tuinman, Blanton, and Gray (1975) obtained evidence which purported that cloze measurements only reflect the influence of local redundancy in contrast with conventional comprehension tests which were uninfluenced by such variables. Carroll (1972) also maintained that cloze scores are dependent mainly on linguistic clues in the immediate environment of the deleted word.

On the other hand, Marshall (1970) studied this problem by using connected discourse, discrete sentence, and fragment levels of contextual constraint and found that subjects could successfully take advantage of increased amounts of contextual information present in connected discourse. Ramanauskas (1972) manipulated sentence order and found that cloze scores within sentences were influenced by constraints operating beyond the sentence level. Some critics have maintained that the procedures used in these studies somehow create artificial conditions for testing the contextual constraint hypotheses. However, Oller (1979) maintained, nonetheless, that such critical evaluations are not valid and that if manipulation of sentence order influences cloze scores within sentences, it cannot, indeed, be maintained that constraints only operate within a few words on either side of a deletion.

The present study stems in part from a statement by MacGinitie (1961) that ". . . although it seems that constraints between words generally decrease very rapidly with distance, this does not mean that constraints never operate over distances of more than four or five words." MacGinitie suggested that knowledge about the topic of a passage might have a more generalized influence that does not decline with decrease in length of context. Therefore, we decided to test the suggestion made by MacGinitie by studying contextual constraints within and between pre- and post-cloze tests as a technique for manipulating "familiarity" as an independent variable. Also, we noted that MacGinitie and others used very easy materials in relation to the reading level of the college age subjects in testing this hypothesis. We believe that constraints might well operate over larger distances for materials of greater difficulty where greater complexity of sentence structure might result in clues further removed from deletions.

The general purpose of this study was to investigate the influence of familiarity upon contextual constraints and gather other evidence related to the construct validity of the cloze procedure. It is important to keep in mind that the construct validity of many reading tests is subject to question, and that this issue does not necessarily have a bearing on the practical use of tests with strong criterion-related validity for many purposes.

PROCEDURE

The experimental procedure for this study involved deleting single words at periodic intervals from a cloze passage and determining how successfully each subject was able to restore the exact words in a cloze test. Subjects were randomly assigned to one of the following two conditions: (1) pre-treatment cloze with no prior exposure to the selection or (2) post-treatment cloze with one prior reading of the passage. Within the pre- or post-treatment cloze condition, the subjects were also randomly assigned to the stimulus passage with one of four omission sets — deletion patterns having every fifth, sixth, tenth, or fifteenth word omitted from the text.

The following experimental hypotheses were tested:

HA_1: Differences in cloze scores will be induced by two levels of familiarity across four deletion ratios.

HA_2: Differences in cloze scores will increase as a function of the magnitude of four deletion ratios across two levels of familiarity.

HA_3: There will be an interaction between the effects of deletion ratio and familiarity level upon cloze test performance.

Two hundred and eighty-four students comprising the 8th grade class of a county junior high school were chosen as subjects. Approximately 20 subjects were randomly assigned to each of the eight experimental conditions in two separate studies. Experimental conditions consisted of combinations of pre- versus post-cloze treatments and four deletion ratios. The two studies constituted partial replication in that each study manipulated the same independent variables using two different sets of words selected from the same passage.

A prose passage entitled, "The Seeing See Little" by Helen Keller was chosen from the 8th level basal of the *Holt Basic Reading System* (Keller, 1977). The 304 word selection was a descriptive narrative. The ability level was estimated to be 8th grade using the Fry readability graph. The mean grade equivalent score for these subjects on the Total Reading Sub-Test of the *Stanford Achievement Test* was 9.0.

Using the same technique as MacGinitie (1960, 1961), the cloze tests were organized so that only those words commonly deleted in all deletion ratios were considered for comparison. This procedure allowed the difficulty of specific words to be held constant while varying the context around these same words. Thus, differences among deletion ratios should represent the effects of context rather than differences in the difficulty of individual words.

Testing was conducted during regular language arts class periods for two consecutive days. Initially, all students were given a booklet, including instructions on taking a cloze test and an appropriate cloze exercise to prepare them for taking the experimental test. Each student took only one test. Students were instructed to supply the one exact word per blank that the author would most likely have used. Students taking the post-cloze tests were given two comprehension questions to answer immediately after reading the article and prior to taking the cloze test. This was done to reduce the effect of immediate memory for specific words upon post-cloze test performance.

RESULTS

The data for both Study I and II (the two studies were designed as partial replications) were analyzed using a two-way analysis of variance. Each study utilized two levels of analysis for the "familiarity factor" and four levels of analysis for the "deletion ratio" factor. Results for Study I indicated significant effects for differences between pre- and post-cloze results and among results for four deletion patterns. No significant interaction was found. On the other hand, analysis of variance results for Study II yielded only one significant *F* ratio — (i.e., pre- versus post-cloze results). Although significant differences were found for Study I in an expected direction favoring the influence of greater context, there was not a consistent rank order of cloze test means from lower to higher deletion ratios. Therefore, results were analyzed by comparing combined deletion ratios for every fifth and every sixth pattern with combined deletion ratios for every tenth and every fifteenth pattern.

Mean percent pre-cloze scores for Study I were 46% for the fifth-sixth deletion ratio and 54% for the ten-fifteenth deletion ratio. Comparable results for post-cloze tests, were 53% and 65%, respectively. The divergence between pre- and post-treatment cloze increased as anticipated with an increase in context, but this interaction was not significant.

For Study II, mean percent pre-cloze results for the fifth-sixth pattern versus the tenth-fifteenth pattern were 41% and 41%, respectively. Comparable deletion pattern results for post-cloze tests were 50% and 46%, respectively. The expected increase in cloze scores between the smaller versus the larger amount of context did not appear as in Study I.

In an attempt to explain the different results from Study I and Study II, several post hoc analyses were carried out. It was observed that the words used in Study I consisted of nine content words and only one function word whereas the words used in Study II consisted of five content and five function words. It is known that function words tend to be more easily predicted in cloze tests than content words, but no studies known to the authors have investigated the influence of familiarity and context upon such word predictions.

Results of the post hoc analysis for Study I and Study II are displayed in Tables 1 and 2.

Table 1
Study I Cloze Test Mean Percent Scores for Content Words as a Function of Levels of Familiarity and Deletion Ratios

	Familiarity	
Deletion Ratio	Pre-Cloze (n=71) Content[1] Words	Post-Cloze (n=72) Content[1] Words
10-15	50.0	63.5
5-6	44.1	51.6
Difference	5.9	11.9

[1] n=9 words

Table 2
Study II Cloze Test Mean Percent Scores for Content and Function Words by Familiarity Levels and Deletion Ratios

	Familiarity			
Deletion Ratio	Pre-Cloze (n=71)		Post-Cloze (n=70)	
	Content[1]	Function[2]	Content[1]	Function[2]
10-15	38.4	44.3	46.5	45.3
5-6	32.4	50.0	43.9	55.0
Difference	6.0	-5.7	2.6	-9.7

[1] n=5 words
[2] n=5 words

For Study I a two-factor ANOVA was performed on the results for content words only. For Study II, a three-factor ANOVA with repeated measures on the third factor was performed.

Since only one function word is contained in Study I, only differences between deletion pattern results for content words were considered. These results are presented in Table 1. Significant differences were found for both the effects of familiarity and greater context upon content word predictions.

Several interesting results were observed in the post hoc analysis of data for Study II. For all words, a significant difference ($p < .03$) was found indicating higher cloze scores for post-cloze than pre-cloze conditions. When considering the effect of deletion ratio upon all words, no significant difference was found. However, in the analysis of function versus content words, a significant difference ($p < .0001$) was found for the higher mean score for function as contrasted with content words. A significant two-way interaction ($p < .05$) was found between word type and familiarity. This interaction revealed that familiarity increases content word inferences more than function word inferences. (In Table II, for deletion ratio 10-15, the pre-post mean difference was 8.1 for content words as contrasted with 1.0 for function words. Similarly, for deletion ration 5-6, the pre-post mean difference was 11.5 for content words and only 5.0 for function words.) Also, a significant two-way interaction ($p < .005$) was found between word type and deletion ratio. This finding reveals that the effect of context is significantly different for content words as contrasted with function words. That is to say, greater context facilitates content word predictions, while greater context inhibits function word predictions. The F values for main effects revealed a significant effect ($p < .003$) for the effect of familiarity on content word inferences but no significant effect of familiarity on function word inferences. No significant F value was obtained for the effect of deletion ratio on content words, but the F ratio approached significance ($p < .06$) for the negative effect of context on function words.

Under Study II post-treatment conditions, the prediction of content words was not facilitated by a larger context. This outcome was inconsistent with the findings of Study I. An analysis of the types of words represented among the nine content words in Study I and the five content words in Study II indicated a more representative sample of different types of content words in Study I. Therefore, it is likely that the content word results of Study I are more valid than comparable results of Study II. Post-cloze function word predictions were less accurately made under conditions of greater context, as was the case under pre-cloze conditions. Also similar to pre-cloze conditions, function words were more easily restored than content words.

Concerning the effects of deletion ratio upon function word predictions, we have a significant interaction between these two variables which allows us to conclude that a difference exists between the facilitative influence of broader context upon content word inferences and the inhibiting influence on function word inference. However, when we consider the influence of context upon function words alone, the difference between the larger and smaller context does not quite reach the conventional criterion of $p < .05$ level of significance. As in the case of content words in Study II, we only are basing our findings on a small sample of five words. A larger sample might yield a more significant difference, particularly in view of the consistency of this finding for both pre- and post-cloze test results even for this small sample.

It is believed that several of these findings are new and of potential significance. However, it is recommended that further research be conducted to replicate these findings. The following section will attempt to provide a theoretical interpretation which will shed some light on the interpretation of these results and their implications for the construct validity of cloze tests as a measure of reading comprehension.

IMPLICATIONS

It is believed that the psycholinguistic interpretation of reading comprehension as developed by Frank Smith (1978) provides an appropriate theoretical frame of reference to interpret and integrate the three major findings of this study and to shed new light on the construct validity of the cloze procedure as a measure of reading. These findings were:

1. Familiarity increases cloze scores on tests with all types of words deleted, but the effect of familiarity is greater upon content words than upon function words.
2. Greater context facilitates cloze inferences for content words, but inhibits cloze inferences for function words.
3. Content words are more difficult to predict than function words.

The third finding is not new, but its importance as evidence for the construct validity of the cloze procedure has not generally been recognized. The first and second results are novel and also important as indications of cloze construct validity. Three principles in Smith's psycholinguistic theory are relevant to the interpretation of these results. First, in reading comprehension redundant information from two sources is used in making inferences; information stored in the nervous system from past experience, and information found on the printed page. Second, two types of predictions are involved in making inferences: focal and global. Third, the difficulty of reducing uncertainty in information processing is related to the number of alternatives which must be considered. Each of these principles will be discussed in relation to the major findings of this investigation.

The finding that the post-cloze testing procedure produces higher scores than the pre-cloze test procedure reflects the relative influence of previously stored information versus information found on the printed page in making responses. This is not a new finding. Taylor (1957) and Rankin (1959) observed this outcome long ago. However, this common finding is of considerable importance in refuting the critical attack upon the cloze procedure as merely a measure of responses to local redundancy clues. Obviously, if local contextual clues were the only stimuli influencing cloze responses, the influence of past experience would not be so strong.

However, the finding that the post-cloze procedure has its effect more upon content than function words is novel, and this interesting outcome is in keeping with several psycholinguistic principles concerning the nature of language. Since function words contain no lexical meaning and are used as one of several parts of a complex system of signals to convey syntactical meaning, it is reasonable to find that the prediction of such words should be more dependent upon contextual clues in the text than upon background experience. Previous exposure to the unmutilated text no doubt enables the reader to acquire both semantic and syntactic information, but all of the semantic meaning is conveyed by content words whereas only part of the syntactic meaning is conveyed by function words. This fact, together with the distinction to be discussed between "global" and "focal" predictions involved in processing information, provides a psycholinguistically tenable interpretation of the differential effect of familiarity upon content versus function word predictions in cloze tests.

The unanticipated and seemingly paradoxical finding that greater context facilitates cloze predictions for content words but inhibits cloze predictions for function words is understandable in terms of the distinction between "global" and "focal" predictions involved in processing information. Global predictions involve general and more long range estimations of what is coming next in a reading passage. These predictions are less concerned with specific details than general outcomes and are formulated more in advance than focal predictions. In contrast, focal predictions arise out of specific events, deal with more local anticipations, and are dispensed with more quickly. It is assumed in this investigation that a parallelism exists between global versus focal predictions on the one hand, and content versus function word predictions on the other. The greater effects of context upon content word predictions are in keeping with the theoretical assumption that content word predictions are more global than focal in nature. Global predictions draw the attention of the reader to broader contextual areas and the cost of making important semantic errors in such predictions should motivate the reader to make use of the increased number of clues available in a broader context. The motivation of the reader should be stronger in making global than in making focal predictions because such inferences involve more far reaching implications related to the purposes for reading.

In contrast, the observation that a larger context actually inhibits the correct prediction of function words has not been made in any prior studies known to the authors and was not anticipated in the planning of this investigation. However, a consideration of these results in the light of the theory of focal versus global predictions, adds a new dimension to our understanding of cloze test results and the reading process. Function words are few in number and occur very often. It is probable that clues to the exact prediction of function words in an immediate area on either side of such deletions, and the reader may not be particularly attentive to such clues because of the availability of other redundant and more potent clues to structural meaning in the remaining context (i.e., word order clues and word ending clues). For these reasons, the reader need not pay careful attention to a larger amount of context in making such predictions. Therefore, the availability of a large amount of contextual information may simply provide more alternative clues than are necessary for making the correct predictions of precise function words. This kind of "information overload" is forced upon the reader and inhibits rather than facilitates exact function word predictions. From a psycholinguistic perspective, it is not surprising, after all, to find that greater context is detrimental to function word predictions.

Conflicting findings in the literature concerning the effects of broader context on cloze test performance may well reflect the use of materials by different investigators which had different percentages of function words in the message. The greater the proportion of function words in a given message, the smaller the influence of greater context for the test as a whole should be. It is important to note that easy materials have a higher percentage of function words than more difficult materials. Previous studies have tended to use easy materials in the study of contextual constraints upon cloze performance. This may well explain some findings to the effect that cloze predictions are not facilitated by words beyond a five word context on either side of the deletion. It is hypothesized that the use of materials with more difficult readability levels and/or the restriction of deletions to content words will expand the boundaries of contextual constraints operating in cloze tests. Confirmation of this hypothesis will support the construct validity of the cloze procedure.

The finding that content words are more difficult to predict than function words was also made by Taylor (1957) and others. This finding is related to the fact that many more content words exist than function words, and therefore more alternatives must be eliminated in order to reduce uncertainty in making predictions for content words. This common finding also has important implications for the construct validity of the cloze procedure as a measurement technique. Both Taylor (1957) and Rankin (1957) found that the prediction of content words is more closely related to knowledge of the topic or subject matter than the prediction of function words or combinations of content and function words. This could hardly be true if the cloze procedure was only a measure of responses to local redundancy clues. Obviously, knowledge of subject matter measured primarily by the prediction of content words is an important source of information in making cloze predictions.

The fact that the cloze procedure is sufficiently sensitive to information processing to produce many results which are congruent with basic principles from a psycholinguistic theory of reading comprehension, provides strong evidence for its construct validity and calls into question the validity of the critical attack upon the cloze procedure as merely a measure of responses to local contextual constraints.

In fact, the whole issue of contextual constraints is not, after all, so important in view of the obvious influence of schemata derived from previous experience in making cloze responses similar to the reading process itself. All reading involves the processing of information in relation to schema. It is postulated that the cloze procedure, far from being merely a test of sensitivity to local contextual clues, may well be considered as a slow motion simulation model for the reading process in general. Unlike conventional comprehension tests, cloze tests measure reading as "information processing" and not merely some hypothetical product as indicated by the ability to subsequently answer a number of questions about semantic and syntactic information which were of no concern to the reader at the time of the reading act per se.

REFERENCES

Aborn, M., Rubenstein, H., & Sterling, T. D. Sources of contextual constraint upon words in sentences. *Journal of Experimental Psychology*, 1959, *57*(3), 171-180.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington: V. H. Winston and Sons, 1972.

Keller, H. The seeing see little. In B. J. Weiss (Ed.), *Great waves breaking*. New York: Holt, 1977.

MacGinitie, W. H. Contextual constraint in English prose. (Doctoral dissertation, Columbia University, 1960). *Dissertation Abstracts International*, 1960, *21*, 126A-127A (University Microfilms No. 60-2020).

MacGinitie, W. H. Contextual constraint in English prose paragraphs. *The Journal of Psychology*, 1961, *51*, 121-130.

Marshall, W. A. Contextual constraint on deaf and hearing children. *American Annals of the Deaf*, 1970, *115*, 682-689.

Oller, J. W. *Language tests at school: A pragmatic approach*. London, England: Longman, 1979.

Ramanauskas, S. Contextual constraints beyond a sentence on cloze responses of mentally retarded children. *American Journal of Mental Deficiency*, 1972, *77*, (3), 338-345.

Rankin, E. F. *An evaluation of the cloze procedure as a technique for measuring reading comprehension*. Doctoral dissertation, University of Michigan, 1957. (University Microfilms, No. 58-3722).

Rankin, E. F. The cloze procedure — its validity and utility. In O. S. Causey & W. Eller (Eds.), *Starting and improving college reading programs.* Eighth Yearbook of the National Reading Conference, Fort Worth: Texas Christian University Press, 1959.

Rankin, E. F. The cloze procedure — a survey of research. In E. L. Thurston & L. E. Hafner (Eds.), *The philosophical and sociological bases of reading.* Fourteenth Yearbook of the National Reading Conference. Milwaukee: The National Reading Conference, 1965.

Smith, F. *Understanding reading.* New York: Holt, 1978.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly,* 1953, *30,* 415-433.

Taylor, W. L. Cloze readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology,* 1957, *41,* 19-26.

Tuinman, J. J., Blanton, W. E., & Gray, G. A note on cloze as a measure of comprehension. *The Journal of Psychology,* 1975, *90,* 159-162.

TAFFY E. RAPHAEL
PETER WINOGRAD
P. DAVID PEARSON
University of Illinois
Center for the Study of Reading

STRATEGIES CHILDREN USE IN ANSWERING QUESTIONS

Research concerning questions has traditionally involved only those areas such as questioning techniques (e.g. Gusnak, 1967), the effects of adjunct questions (e.g. Anderson & Biddle, 1975), the development of question taxonomies (e.g. Barrett, 1976; Bloom, 1956), and children's performance on different types of questions (e.g. Bormuth, Manning, Carr, & Pearson, 1970). While related to the question taxonomy literature, this study looks at questions in a heavily context dependent fashion; that is, the basic thesis of the study is that in order to understand the impact of a question upon a student's comprehension, one must consider: (1) the relationship between the kind of question asked and the kind of response given, (2) the source of the information available for responding, i.e. the text or the reader's base of prior knowledge, and (3) the student's awareness of these complex relations between the question, response, and knowledge source.

Comprehension questions are the major means for assessing readers' understanding of text; they are found in standardized tests, informal reading inventories, and content area texts to name only a few. In short, it is performance on comprehension questions which often determines whether or not a child is considered to be a successful student or whether he/she is labeled a "poor reader." Therefore, one means for facilitating children's comprehension, at least as it is operationalized in a school setting, may be to improve their ability to perform on the questions.

A line of research that could prove fruitful in developing instructional strategies to help the poor reader is to examine some factors that may influence children's ability to respond to comprehension questions, and determine how these factors differentiate between the good and poor readers. Reviews of the research (Golinkoff, 1975-76; Ryan, in press) concerning comprehension problems suggest that some of the differences between good and poor readers may lie in their selection and use of appropriate strategies. Strategy differences between good and poor readers in answering post-reading comprehension questions is a potential area for investigation.

To investigate the use of question-answering strategies, it is relevant to study behavior as readers answer different types of questions, examining the kinds of responses given to various types of questions. The literature in the area of taxonomies for classifying the types of questions seems abundant, particularly if related to general taxonomies of comprehension (Bloom, 1956; Barrett, 1976), yet rarely do these taxonomies consider the text to which a question refers. The same question may vary in its type of classification depending upon the passage to which it refers, literal in a text where the answer is explicitly stated, inferential in the case where the answer is not stated in the text. Pearson and Johnson (1978) propose a taxonomy which takes into consideration the relationship between questions and text. Their three category taxonomy provides for the relationship between the information stated in the text and the information that must come from the readers' knowledge base. Questions and responses are classified as *text explicit* when the question invites a response that is explicitly stated in the text, *text implicit* when the question invites the reader to draw a text-based inference which could involve the integration of information across sentences or paragraphs, or as *script implicit* when the question invites inferences that require the reader to integrate information in the text with information in his or her store of prior experience. (See Table 1 for examples).

We stress the concept of "invited response" for even though the question can be categorized by the type of response it is *likely* to elicit, it does not necessarily require or demand that response. The relationship can be described as unidirectional in that one can depart from the explicit message and answer a text explicit question with knowledge-based information, but one cannot return to the text to locate an answer to a script implicit question since the information would not be there.

To summarize, the purpose of this study is to investigate the performance differences between good and poor readers as they respond to post-reading comprehension questions, focusing on strategy-selection in answering three categories of these questions: text explicit, text implicit, and script implicit.

METHOD

Subjects

Two hundred and forty fourth, sixth, and eighth grade students from a suburban school system northwest of Chicago participated in the study, 80 at each grade level. Fourth grade students were drawn from two different elementary schools while the sixth and eighth grade students attended the same junior high school. Students were classified as either high or low ability readers on the basis of a combination of reading comprehension scores on the Iowa Test of Basic Skills and the developmental reading group to which they were assigned.

Design

A 2^5 X 3 factor design was used for this study to assess the effects of familiarity (more or less), task order (read first or rate first), reading ability (high or low), text access (access or no access), bias (biased or unbiased), and question type (text explicit, text implicit, script implicit) on children's ability to predict how they would answer a given question and on their actual performance in answering the questions.

Task order, text access, and reading ability were between-subjects factors in all grade levels, as was familiarity at the eighth grade only. Within-subject factors were question type and bias in the three grade levels and familiarity at the fourth and sixth only.

Six dependent measures were used:

1. Total Number of Comprehension Questions Answered Correctly
2. Total Number of Text-Based Responses

3. Number of Text-Based Responses Answered Correctly
4. Total Number of Script-Based Responses
5. Number of Script-Based Responses Answered Correctly
6. Number of Questions Correctly Rated as Text Explicit, Text Implicit, or Script Implicit

The first five measures involved the students' actual performance on the comprehension questions, and the final measure concerned the students' predictions about how they would answer the questions.

Materials

At each grade level, experimental materials consisted of five items: (1) two passages, one familiar and one unfamiliar (based on independent ratings by adult judges), (2) a list of the eighteen comprehension questions with a categorizing task to classify each question according to its predicted type of relationship with the text [text explicit (operationalized as RIGHT THERE), text implicit (operationalized as THINK AND SEARCH), and script implicit (operationalized as ON MY OWN)], (3) a list of the eighteen questions with space provided for answer, (4) directions for administration, and (5) a series of interpolated tasks to allow short term memory to decay.

Passages varied in length and readability in order to be appropriate for the grade level at which they would be used. They ranged in length from 400 words at the fourth grade to 800 words at the eighth grade. Readability was determined by use of the Fry Readability Formula.

The comprehension questions were equally divided into either text explicit, text implicit, or script implicit types, with two forms of each question developed to measure "bias." To create a biased question, the words "do you think" were embedded within the question as indicated by the following example:

(1) What was the main job of the spring round-up? (UNBIASED)

(2) What do you think was the main job of the spring round-up? (BIASED)

For the list of eighteen comprehension questions, two different random orders of questions were generated. The purpose of the first was to randomize the inclusion of the biased versus unbiased question, and the second was to randomize the order of presentation of the questions. Thus, a total of eight experimental conditions were necessary.

Eight different experimental booklets were created at each grade level, one per condition, each consisting of nine sections. All booklets began with an explanation of the experimental task, a description of the three types of questions, and a sample passage followed by one of each question type. The second and third sections were composed of either a passage followed by the prediction task, or vice versa, and the fourth contained comprehension questions for the students to answer. Following the passage and the comprehension task were interpolated activities which consisted of word or number searches. The same order of activities was used for the second passage. Following both passages and activities was a twelve item debriefing questionnaire.

Procedure

The procedure varied slightly at each grade level because of the facilities available in the respective schools and scheduling constraints for the different grades. Student booklets were similar (containing passages appropriate for the grade level and balancing for order of the experimental task), and all testing was done over two sessions, one passage per session.

Sixth and eighth grade students were tested in four groups of twenty in an open classroom area. Fourth graders were tested in groups of ten in individual classrooms.

After each child received a test booklet, the purpose of the experiment was explained. Following a description of the three question types, students did a sample exercise consisting of a short story and one each of the three types of questions. The third and fourth task varied across groups: either they read the story silently, then rated the questions, or vice versa. Questions for rating were read aloud by one of four experimenters. The fifth step involved answering the debriefing questions, again read aloud by the experimenters.

TABLE I

Response Scoring

Categorization:	**Broad**	**Strict**	
TEXT	10		Any and all text-based responses
Total Deleted Constituent		11	Response drawn word-for-word from the text
Partial Deleted Constituent		12	Word-for-word from the text, but some part of the required response is deleted
Correct Text Integration		13	Reasonable answer from the text, but not deleted constituent from the sentence from which the question was constructed
Incorrect Text Integration		14	Formed erroneous conclusion from linking material from the text inappropriately
Text Dependent		15	Copied irrelevant material from the text with no apparent regard for content of the question
Question Dependent		16	Response drawn from information or words in the question
Causal Linking		17	Response can be traced as part of a chain of events leading to or from the deleted constituent
ANAPHORA	20		Any synonymous response: can be pronomial or lexical in nature
Superordinate		21	Higher order, higher level, more general response
Coordinate		22	Synonymous, at the same level of the deleted constituent
Subordinate		23	Subset of the required response
SCRIPT	30		Any and all knowledge-based responses
Passage Related		31	Evidence of having read the story present
Passage Independent		32	Sensible answer, but could be given without knowledge of the passage
Inappropriate/Wrong		33	Inference that is totally unreasonable, wrong
UNSCORABLE	40		Cannot assign one of the above
Blank		41	No answer or "I don't know"
Illegible		42	Incoherent or not interpretable

TABLE 2
Sample Scoring From Fourth Grade Passage

They never stay in one place very long, because they must always look for food and water. So they never build houses. Most of the time, aborigines wear no clothes. If it's very cold, they wrap animal skins around themselves.

Aborigines travel together in small groups. Whenever they find a place where there is water or a few trees, they set up a camp. But they don't have tents. Each family makes a screen of tree branches or bark. This protects them from wind and bad weather.

Question type	Question	Score	Sample Response
TE	When and where do Aborigines set up camp?	11	whenever they find a place where there is food and water.
		12	where there is water
		13	where they think they can hunt animals (note: from elsewhere in the text)
		15	tree branches or bark
		22	near a stream to get water, they set camp near trees
		23	where they have water
		32	in a cave
		33	at noon
TI	What protects the Aborigines from wind and bad weather?	14	they wear animal skins to keep them warm
		17	building a shelter
		21	shelters from products of trees
SI	Why do Aborigines travel in small groups instead of big groups?	16	they move in small groups
		31	so they won't have to hunt a lot of food
		42	the chiterans and wilfs you to go by
		41	I don't know

The procedure was then repeated at the next session for the other passage for fourth and sixth grade students only. Interpolated tasks were used after each of the reading tasks and the comprehension questions to control for the effect of short term memory as well as to allow the slower students time to catch up.

Response Scoring

The scoring system was an expansion of Jackson's (1972) and Pearson and Nicholson's (1976) work on children's responses to comprehension questions. Interrater reliability using the system was .94 overall (.95 in fourth grade, .96 in sixth, and .92 in eighth).

Based upon previous research as well as children's actual responses the system was designed to operate under either a broad or a strict criterion. The broad criterion system consists of four categories of responses: (1) text (the answer was from text-based information), (2) anaphora (the answer was synonymous to or an anaphorization of information in the text), (3) script (the answer was derived from the reader's knowledge base), and (4) unscorable. The broad score is represented by the ten's digit in the numerical code and the strict scoring system which subdivides the four broad categories is represented numerically by the one's digit (see Table 1). For example, all anaphoric responses would be coded as 20, using the broad criterion only. Using the strict, a superordinate anaphora is coded as a 21, coordinate as 22, and subordinate as 23. A more complete description of categories with definitions and examples can be found in Tables 1 and 2.

Results and Conclusions

Three issues were investigated in this study: (1) children's performance in responding to comprehension questions, (2) their awareness of task demands of different types of questions, and (3) a taxonomy of questions and responses based on the relationship between questions and text. A summary of relevant significant effects for the five measures of children's responses to the questions can be seen in Table 3. (A more complete version of this paper, including further analysis, is available on request.)

Across all measures at all grade levels, results give strong evidence to support the suggestion that question type does influence the type of response given. Children behaved predictably such that more text than knowledge-based responses were given to text-based questions, and more knowledge than text-based responses were given to the knowledge-based questions. In addition, predictable significant effects for ability were present across most of the measures (see Table 4).

TABLE 3

Summary of Significant Main Effects
and Two Way Interactions:
Response to Questions

Main Effects	Measure				
	Correct Responses	Text Responses	Correct Text Responses	Script Responses	Correct Script Responses
Ability (A)	4, 6, 8*	4, 6, 8	4, 6, 8	4	4, 6, 8
Task (T)	8	8	8		
Access (Ac)			6		
Bias (B)					
Familiarity (F)	6	4	4, 6	4, 6	4, 6
Question Type (Q)	4, 6, 8	4, 6, 8	4, 6, 8	4, 8	4, 8
A x Q		4, 6	4, 6, 8	4, 6	4, 6, 8
T x Q	4		8	8	8
Ac x Q	6				
F x Q	4	4, 6	4, 6	4, 6	4, 6

*Numbers refer to grade level.

The second issue, the awareness of task demands of questions, was examined by analyzing children's ability to correctly identify a question as text explicit, text implicit, or script implicit depending upon the type of response it is likely to elicit. Results can be seen in Table 5.

TABLE 4
Main Effects for Ability

Grade	Fourth		Sixth		Eighth	
Ability	**High**	**Low**	**High**	**Low**	**High**	**Low**
Measure						
Correct Responses	2.50	1.79	2.24	1.25	1.96	1.31
Textual Responses	1.55	1.22	1.73	1.12	1.44	1.07
Correct Textual Responses	1.35	.80	1.42	.74	1.13	.71
Script Responses	1.35	1.54	- -	- -	- -	- -
Correct Script Responses	1.14	.98	.81	.51	.83	.60

TABLE 5
Question Rating Task

	Question Type		
	Text Explicit	**Text Implicit**	**Script Implicit**
Grade			
4	2.338*	2.089	2.46*
6	2.298*	2.00	1.92
8	2.215	2.658*	2.848*

*Significantly better than chance $p < .01$

Looking across grade levels, children appear to be sensitive to the differences between the three types of questions. It is difficult to speculate about developmental trends since children at each grade level answered questions drawn from different passages. In a second analysis, when ratings of and responses to questions were compared, the data suggest that good readers tend to behave in a more internally consistent manner than do poor readers. That is, they tend to give responses consistent with their predictions. What remains to be determined is the exact nature of the relationship between task awareness and task performance.

The question type variable elicited consistent and predictable main and interaction effects and the response scheme proved to be objective across three independent judges.

In summary, we examined children's performance on post-reading comprehension questions as a function of question type. Specifically, we were interested in whether or not they recognized differences in question type and how this influenced their performance. While we cannot be unequivocal in our conclusions, at least it appears that children can distinguish between questions and that question type definitely affects the response given. What remains to be studied is whether sensitivity to question type affects performance on comprehension questions. The relationship between general comprehension of the text and performance on comprehension questions also needs to be determined.

REFERENCES

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York: Academic Press, Inc., 1975.

Barrett, T. C. Taxonomy of reading comprehension. In R. Smith & T. C. Barrett (Eds.), *Teaching reading in the middle grades.* Reading, Mass.: Addison-Wesley, 1976.

Bloom, B. (Ed.), *Taxonomy of educational objectives.* New York: McKay, 1956.

Bormuth, J. R., Carr, J., Manning, J., & Pearson, P. D. Children's comprehension of between- and within-sentence syntactic structure. *Journal of Educational Psychology,* 1970, *61,* 349-357.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly,* 1975-76, *11,* 623-659.

Guszak, F. J. *Relationship between teacher practice and knowledge of reading theory in selected classes* (Project Report Number 5-437). Washington, D. C.: United States Department of Health, Education, & Welfare, 1966. (ERIC Document Reproducton Service No. ED 010 191)

Jackson, L. A. *A schema for classifying and scoring student constructed literal, anaphoric, and inferential responses to literal comprehension test items.* Unpublished masters thesis, University of Minnesota, 1972.

Pearson, P. D. & Johnson, D. D. *Teaching reading comprehension.* New York: Holt, Rinehart, & Winston, 1978.

Pearson, P. D. & Nicholson, T. *Scripts, texts, and questions.* Symposium paper presented at the National Reading Conference, Atlanta, November 1976.

Ryan, E. B. Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), *Advances in reading research* (Vol. 2). New York: Academic Press, in press.

BARBARA TAYLOR
SANDRA BERKOWITZ
University of Minnesota

FACILITATING CHILDREN'S COMPREHENSION OF CONTENT MATERIAL

Elementary teachers frequently comment that the content area textbooks that they are expected to use in the classroom are too difficult for many of their students to read. However, elementary students in the intermediate grades must learn to comprehend and remember what they read in their content area textbooks because these texts are important components of the elementary curriculum.

Research on study techniques to facilitate elementary students' comprehension and memory of content area material is limited. Most of the work in this area has been directed toward secondary and college students (Anderson, in press). Therefore, the purpose of the present study was to begin to fill this gap by examining the relative effectiveness of three study techniques on elementary students' comprehension and memory of social studies material. The three study techniques under investigation include the following: generating summary statements for paragraphs while reading, completing a study guide, and answering questions after completing a reading assignment.

The first study technique of interest was the summary generation task. Students using this technique generated a one-sentence summary for each paragraph that they read before going on to the next paragraph. Current research suggests that techniques that require students to generate responses while reading tend to facilitate comprehension of the material (Frase & Schwartz, 1975; Ross & DiVesta, 1976). Andre and Anderson (1978-79) found that high school students who generated questions for paragraphs as they read performed significantly better on a comprehensive test than students who read and reread the same material. Also, Doctorow, Marks, and Wittrock (1978) found that sixth grade students' reading comprehension was enhanced when they wrote a one-sentence summary after each paragraph read. Based on these findings, it was hypothesized that the summary generation procedure, in comparison with the other study techniques under investigation, would facilitate sixth grade students' comprehension and memory of material from a social studies textbook.

The second technique of interest involved the completion of a study guide. The study guide directed the students to read the text one page at a time and to stop to answer 3-4 questions at the end of the page before reading further. While little research exists on the effectiveness of study guides per se (Patberg, 1979), adjunct questions following short segments of text have been found to facilitate reading comprehension (Anderson & Biddle, 1975; Rothkopf & Bisbicos, 1967). Based on these findings, it was hypothesized that the study guide procedure, in comparison to the conventional procedure and the control procedure, would enhance sixth grade students' comprehension and memory for content material.

The third study technique was a post questioning procedure. Students read a five page text segment in its entirety and then responded to questions over the material. This procedure was included as a conventional classroom study technique. A control condition was also included in the investigation in which students read a social studies passage and then performed a distractor task unrelated to the content of the material.

METHOD

Subjects

One hundred and eight sixth grade students reading on or above grade level from a suburban elementary school served as subjects. Reading level was determined by teacher judgment and placement in basal reading material. Students were randomly assigned to one of the three study groups or to the control condition.

Materials

All students read a 700-word passage from a contemporary sixth grade social studies textbook. A set of post questions and a study guide were prepared for this material. Both study aids contain the same questions.

Procedure

All subjects participated in a one hour training session in which they practiced their assigned study technique on a social studies passage similar in length to the test passage. The following week the students read the test passage and performed their assigned study task. Each group was instructed to read over their completed assignment. Then, they took a short answer comprehension test on the material. The next day the students wrote a recall of everything they could remember from the test passage and completed a multiple choice test on the material. The short answer and multiple choice tests were comprised of new questions that probed information that had not been tapped by the study guide or post questions.

A discourse analysis procedure adapted from the work of Grimes (1975) and Meyer (1975) was used to segment the test passage into propositions and place the propositions into a hierarchical structure of main ideas and details. The students' written recalls were divided into propositions and scored against the passage template for total number of propositions recalled and number of main idea propositions recalled. A random sample of 16 recalls were scored by a second scorer to establish inter-rater reliability. The Pearson *r* was .94 for the two sets of total recall scores and fo the two sets of main idea scores. The short answer tests and multiple choice tests were scored against a scoring key for number of items correct.

RESULTS

The means and standard deviations for each group on each dependen measure are presented in Table 1.

First, a multivariate analysis of variance was performed on the data to compare the four groups on the four dependent measures: short answe score, multiple choice score, total recall score, and main idea score. The results indicated that there was at least one difference among groups on

TABLE 1
Mean Scores and Standard Deviations
for the Four Groups on the Four Dependent Measures.*

		Score			
Group		Multiple Choice (total possible = 20)	Short Answer (total possible = 20)	Total Recall (total possible = 85)	Main Idea Recall (total possible = 30)
Summary	$\bar{X}$	15.26	12.56	10.07	5.60
	s.d.	2.55	3.90	5.22	3.54
Study Guide	$\bar{X}$	15.63	13.00	6.74	2.70
	s.d.	2.08	2.95	3.06	1.90
Post Questions	$\bar{X}$	15.33	11.74	7.59	3.41
	s.d.	2.30	2.65	3.31	2.15
Control	$\bar{X}$	14.37	11.81	6.63	3.07
	s.d.	3.20	3.04	2.91	2.02

*n = 27

their composite comprehension measure, approximate $F(12,262) = 2.20$, $p < .05$.

Univariate analyses of variance were then used to compare the four groups on each of the dependent measures. No differences were found among groups on the short answer or multiple choice scores ($p > .05$). However, it was found that there was at least one difference among groups in terms of the total recall score, $F(3,10) = 4.95$, $p < .01$; and in terms of the main idea score, $F(3,10) = 7.31$, $p < .001$. Newman-Keuls post hoc tests revealed that the Summary Group recalled more propositions in general and more main idea propositions than the other three groups ($p < .05$). No other differences among groups were found.

Students in the Summary Group may have recalled more main idea propositions than the other three groups simply because they recalled more propositions in general. To examine the four groups' recall of main ideas more closely, t-tests for correlated data were used to compare the proportion of main ideas to details recalled by each student with the proportion of main ideas to details in the test passage. This analysis revealed that only the Summary Group recalled more main ideas than expected, based on the proportion of main ideas to details in the passage, $t(26) = 4.84$, $p < .001$.

DISCUSSION

The purpose of this study was to investigate the effectiveness of several study techniques on sixth grade students' comprehension and memory for content area textbook material. One technique of interest, the study guide procedure, received no empirical support as a particularly effective study technique. Students who used a study guide while reading a social studies passage did not perform significantly better on measures of comprehension and memory over this material than students who answered questions after reading the entire passage or students who simply read the passage.

The other technique of interest, the summary generation procedure, did appear to enhance students' memory in general for the social studies material that they read. Additionally, the procedure appeared to improve their memory for the main ideas of the material. The data analysis revealed that the students in the Summary Group recalled more propositions in general and more main idea propositions than the students in the other three groups under investigation. These findings suggest that writing a one-sentence summary after each paragraph while reading content area material may be a valuable study technique for elementary students. While additional research is needed, the findings of this study also support the notion that memory of expository material is enhanced when students generate responses to the material while reading as opposed to simply responding to the material in a more passive manner such as answering questions.

Elementary teachers may find the summary generation technique to be a desirable classroom tool because of several additional factors. First, it appears that the technique may require only a minimal amount of training, as the students in this study practiced the procedure for just one hour before being tested. Second, students can use the summary generation procedure without teacher input. Therefore, teachers do not have to spend preparation time constructing study questions when this technique is used.

More importantly, students are gaining skill in an independent study technique when practicing the summary generation procedure.

REFERENCES

Anderson, R. C., & Biddle, B. On asking people questions about what they are reading. In G. Bower (Ed.), *The Psychology of Learning and Motivation*, 1975, *9*, 89-132.

Anderson, T. Study strategies and adjunct aids. In R. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: L. Erlbaum, in press.

Andre, M. & Anderson, T. The development and evaluation of self-questioning study technique. *Reading Research Quarterly*, 1978-79, *14*, 606-623.

Doctorow, M., Marks, C. & Wittrock, M. Generative processes in reading comprehension. *Journal of Educational Psychology*, 1978, *70*, 109-118.

Frase, L. & Schwartz, B. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 1975, *67*, 628-635.

Grimes, J. *The thread of discourse*. The Hargue, Netherlands: Mouton and Company, 1975.

Meyer, B. *Organization of prose and its effects on recall*. Amsterdam: North Holland Publishing Company, 1975.

Patberg, J. Validation of reading strategies in secondary content areas. *Journal of Reading*, 1979, *22*, 332-336.

Ross, S. & Divesta, F. Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology*, 1976, *68*, 689-695.

Rothkopf, E. & Bisbicos, E. Selective facilitative effects of interspersed questions on learning from written material. *Journal of Educational Psychology*, 1967, *58*, 56-61.

SUSAN THOMAS
CONNIE BRIDGE
University of Kentucky

A COMPARISON OF SUBJECTS' CLOZE SCORES AND THEIR ABILITY TO EMPLOY MACROSTRUCTURE OPERATIONS IN THE GENERATION OF SUMMARIES

The purpose of this study was to examine the relationship between subjects' cloze scores and their ability to employ macrostructure operations in the generation of summaries. Previous studies by MacGinitie (1960, 1961) and Carroll (1971) have indicated that cloze test scores reveal mainly sensitivity to local redundancy clues within the immediate context on either side of word deletions. Therefore, it has been argued that cloze tests do not measure high level cognitive processes involved in reading comprehension. On the other hand, studies by Thomas (1980), Darnell (1963), Ramanauskas (1972), and others have yielded evidence which suggest that contextual constraints in cloze passages may operate over a more extensive area than that found by previously mentioned studies. The investigation by Thomas revealed that contextual constraints within and between sentences do operate over longer distances than suggested by MacGinitie and Carroll. Also studies by Darnell and Ramanauskas have shown that scrambling sentence order has a significant, negative influence on cloze performance. Thus, it was postulated that the cloze task might be affected by some of the same underlying factors that influence comprehension at the macro-structure level as measured by a recall task. Some of these factors will be discussed in relation to the characteristics of subjects' summaries.

Recalls have been used as a measure of comprehension based upon the assumption that the information rendered in the recall situation is a reflection of the knowledge structure of the reader after having processed the information in the text. Recalls include both information that was explicitly stated in the text plus information that was generated or inferred by the interaction of the reader's knowledge of language and the world with the printed message of the author. When asked to produce a recall or summary of the text, subjects' recalls reveal that during reading and recall various cognitive operations have been performed upon the propositions in the original text to arrive at the propositions in their recalls. These operations involve the *integration* of several propositional units into a single macro-proposition, the *generalization* of several specific ideas into a more generic idea, and/or the *deletion* of information that subjects deem non-essential (Kintsch 1978; vanDijk 1977).

PROCEDURE

Eighteen students were randomly selected from an eighth grade class of a Fayette County, Kentucky, junior high school to participate in this study. The students were asked to first complete a cloze task with no prior reading of the intact passage. Secondly, the students were allowed to read the unmutilated text and then required to write a summarization of ideas recalled from the passage.

The passage used in the study was a 304 word expository selection entitled "The Seeing See Little" by Helen Keller which was taken from Level 8 of the *Holt Basic Reading Systems*. The readability level of the passage was eighth as measured by the Fry Graph. The passage presented Keller's opinion that people who possess sight and hearing rarely appreciate these abilities or make the fullest use of them. Only when lost are these faculties appreciated. A personal anecdote about a sighted friend's lack of awareness was included by Keller to support her opinion. Students were not told the title or author at any point in the procedure. An every fifth word deletion pattern was chosen so that the results of the study could be discussed in relation to other cloze studies, since the every fifth deletion pattern is the most frequently used.

Procedure for Scoring Summarizations

A system of scoring the summarizations was developed based upon the assumption that a good summary should reflect the most important ideas in the text. A method based upon that of Johnson (1970) was used to identify the discrete idea units within the text and to rate the structural importance of these units. Raters were asked to read the intact text and to mark the points which they believed to represent divisions between idea units, which were defined as units containing a complete thought. Seventy-five percent agreement by the raters was used as the criteria for dividing the text into units. The twenty idea units thus identified were then typed as a list of statements with each unit beginning on a new line. Graduate students in a reading education class were then asked to rate the importance of these idea units by first reading the passage through once in its entirety and then marking out the five least important ideas, labeling them with a 1. Then they were to go back through the 15 remaining idea units in the passage and mark out the five least important ones, labeling them with a 2. Next they were to read the ten idea units which were left, crossing out the 5 of lesser importance, labeling them as 3's. The five remaining units should then have represented the units judged to be most important; they were treated as 4's. After tallying the ratings, the five units which had received the highest number of points were judged to be the most important ideas within the text, based upon the fact that they had been ranked high by a large number of raters. Each idea unit was accordingly assigned a rank of 4, 3, 2, or 1 with 4 indicating a high importance rating.

The written summarizations generated by the subjects were divided into idea units by the experimenters. Each unit was then matched with its counterpart in the original text when possible. A scale of 0-3 was used to rate the degree to which the recalled unit preserved the meaning of the original unit. It was rated as a three if it was a verbatim recall or a good paraphrase of the original unit. It was rated as a two if it was a verbatim recall or a good paraphrase of a major part of the unit. It was rated as a one if it was a somewhat vague paraphrase or consisted only of a small fragment of the original unit. It was scored as a 0 if it could not be matched to any unit within the text, such as when readers generated incorrect inferences, generated a conclusion that was a logical possibility but not stated in the text, or generated text external information which conveyed no text-related information.

The score for each unit in the summarization was then weighted on two factors: the rated structural importance of the original unit with which the recalled unit had been matched and the degree to which the recalled unit preserved the meaning of the original unit with which it was matched. Thus, a verbatim recall or a good paraphrase of one of the units rated highest in importance, that is, a unit ranked as a 4 would be worth 12 points in the summarization score [4 (structural importance of original unit) x 3 (level of meaning retention in recalled unit)].

RESULTS

A comparison of subject's cloze scores with their summarization scores was made by using a Pearson's product moment procedure and revealed a high correlation (*r*.80) between subject's scores on the cloze test and their ability to employ macrostructure operations in the generation of summaries.

Comparison of the Summarizations of High and Low Cloze Scorers

To look at the differences between the high and low cloze scorers, the subjects were divided into two equal groups based upon whether they scored above or below the median score on the cloze task. Their summarizations were then examined in regard to number and completeness of idea units recalled and the degree to which the recalled units retained the meaning of the original idea units in the passage. This comparison of the number and saliency of idea units recalled and the degree of meaning retention in the recalled units revealed the following differences in ability to generate summarizations.

The students in the high cloze group included more than twice as many idea units in their summaries as did the students in the low cloze group. The mean number of units recalled by high cloze scorers was 3.77 while the mean number of units recalled by low cloze scorers was 1.77.

The students in the high cloze group recalled more units which were rated of highest structural importance than did students in the low cloze group. Of the units rated 4, indicating high structural importance, high cloze scorers recalled a mean number of 1.55 units, while the mean units recalled by poor cloze scorers was only .44. Of the units rated 3, the mean number of units recalled by high cloze scorers was 1.44 and by low cloze scorers was .88. Very few units rated low in importance were recalled by either high or low cloze scorers.

TABLE I

Number of Idea Units at each Level
of Importance Recalled by High and
Low Cloze Scorers

Importance Rating	LoC	HiC
4	4	14
3	8	13
2	3	2
1	1	2

Note: Rating of 4 indicates highest importance.

The pattern of units recalled by high and low cloze scorers was very similar in that both groups tended to include the same units in their recalls and to delete the same ones. Of the ten units rated 2 or 1, indicating low importance, five were not recalled at all, while four others were only recalled by one subject. The only major difference in the pattern of recall between the high and low cloze scorers was in the superior recall of rank 4 idea units by the high group.

The tendency of the low cloze scorers to fail to recall several of the level 4 units which contained more specific concepts, but instead to recall generalized or integrative macro-propositions, concurs with a finding in previous studies of inferred information in the recalls of good and poor third grade readers (Bridge, 1977; Bridge, Tierney, & Cera, 1978; Tierney, Bridge, & Cera, 1979). The poor readers tended to overgeneralize and recall superordinate concepts rather than the specific concepts in the original passage; thus their recalls lacked the precision and specificity of the good reader's recalls.

There was also a difference between the degree to which the high and low cloze scorers' preserved the meaning of the original textual units.

TABLE 2

Number of Idea Units Recalled at
Three Levels of Meaning Retention
by High and Low Cloze Scorers

Meaning Retention Rating	LoC	HiC
3	1	3
2	7	24
1	8	5
0	6	2

Only 36% (8 of 22 units) were recalled by low cloze scorers with enough veridicality and completeness to be rated as a good paraphrase of a major part of the unit, while 79% (27 of 34) of the units recalled by high cloze scorers received a rating of 2 or 3, indicating high meaning retention and completeness. Thirty-six percent (8 of 22) of the units recalled by low cloze scorers were rated as 1 in meaning retention, indicating a vague paraphrase or the recall of an isolated fragment of the unit. Only 15% of the recalled units of high cloze scorers fell into this low meaning retention category. The low cloze seventh grade subjects in this study resembled the poor third grade readers in earlier studies by Tierney, Bridge, and Cera (1979) and the non-fluent adult readers in a study by Marshall (1976) in their tendency to recall fewer propositions and less complete propositions.

Low cloze scorers also generated more propositions which could not be matched with any unit in the original text; 27% (6 of 22) of the recalled units were of this nature. Five of that six were vaguely related to the notions of blindness and deafness but were so vague as to be unscorable; one of the six was simply an erroneous conclusion drawn from the text information. Only 6% (2 of 34) of the recalled units of high cloze scorers represented overgeneralized information. The tendency of the low cloze scorers to recall overgeneralized information also agreed with the characteristics of poor third graders in previous studies (Bridge, 1977).

Analysis of Idea Units Recalled in the Summaries

The most surprising finding was not found in the comparison between high and low cloze scorers, but in the comparison of the units which were most frequently included in all of the summaries with the units which had been ranked highest in structural importance in the preliminary phases of the study. The experimenters had predicted that the level 4 units would appear most frequently in the recalls, with the level 3 units being recalled with the next highest frequency. In actuality, the units most frequently recalled ranked as shown in Table 3.

TABLE 3

Importance Ratings of Most Frequently Recalled Idea Units

Unit Number	Frequency of Recall			Importance Rating
	T	HiC	LoC	
17.	10	7	3	3
6.	8	4	4	3
2.	6	6	0	4
1.	5	5	0	4
5.	5	3	2	2
7.	4	1	3	4
4.	3	2	1	4
16.	2	1	1	3
10.	1	1	0	1
11.	1	0	1	1
15.	1	1	0	3
19.	1	1	0	1
20.	1	0	1	2

In an attempt to explain these unexpected results, it was necessary to compare the two tasks that readers performed on the text: the rating task performed in the preliminary stage of the study as well as the summarization task itself. It was also necessary to look at characteristics within the text to understand why specific lower ranked units appeared in the summaries more frequently than some of the highest ranked level 4 units.

It is hypothesized that the rating of the structural importance of the twenty idea units when the entire text was available in the form of a sequential list of idea units required the raters to deal with text at a microstructure level. (A copy of the text parsed into idea units appears in Appendix 1). Indeed Kintsch (1978) has labeled the list of the propositions stored within a subjects' memory as the propositional microstructure. In the rating task the subjects had a sequential list of all of the idea units to peruse; thus, it was possible to deal with each individual unit in its explicitly stated form, weighing the relative importance of each to the understanding of the entire text. On the other hand, the production of a summary from memory reflects the macrostructure or gist of the text. Kintsch and vanDijk believe that when readers process text, the representation of the meaning of the text in the subject's memory consists of the propositional microstructure, which is organized into larger units which have macrostructure propositions as labels. Macro-propositions are inferred from or identified in a text base and serve to summarize the text in memory.

The mental operations employed in the generation of macro-propositions include integration, deletion, and generalization, (vanDijk, 1977; Kintsch, 1978). The integration macro-rule was determined to be operating within the target text in this study, as the integration rule states that a proposition may be identified as a summarizing unit if the other propositions in the text base are normal conditions, components, or consequences of it. Idea unit 5 which was rated as a 2, "Their eyes and ears take in all sights and sounds hazily, without concentration, and with little appreciation," serves to summarize idea units 1, 2, 3, & 4, three of which were rated as level 4 in structural importance.

The integration function of unit 5 can be further explicated by an analysis of the cohesive ties which relate it to the preceding units. The anaphoric pronoun "Their" refers back to unit 3 to "those who have never suffered impairment of sight or hearing." "Their" is an identifying reference item and thus functions as a deictic and is specific (Halliday & Hasan, 1976). "Their" and "Those" have an identity relationship. The words "eyes" and "ears" and "sights" and "sounds" represent a type of lexical cohesion known as collocation, which results from the mutual expectancy that certain words will occur frequently in the same situational context. The use of these words in unit 5 serves to provide cohesive ties with unit 1 through the words "deaf" and "hearing", with unit 2 through "blind" and "sight", with unit 3 through "sight" and "hearing." Additional cohesive ties between unit 5 and units 1 and 2 were provided through lexical cohesion of a reiterative nature by the repetition of the word "appreciation" which appeared as "appreciate" in unit 1 and which is roughly synonymous with the phrase "realize the manifold blessings" in unit 2. Finally, the words "hazily" and "without concentration" in unit 5 appear to be reiterative of the phrase "seldom make the fullest use" in unit 4. Due to the integrative function served by unit 5 and the numerous cohesive ties with the 4 preceding idea units it is not surprising that it was one of the most frequently recalled idea units even though it was only given an importance ranking of 2 by the raters, who tended to rank the more specific units 1, 2, and 4 as higher in importance.

Unit 6 was the second most frequently recalled idea unit, although it only had an importance rating of 3. Applying the Kintsch and van Dijk macro-rules reveals that unit 6 is arrived at through the mental operation of generalization. "It is the same old story of not being grateful for what we have until we lose it, of not being conscious of health until we are ill" is a statement at a generalized level of the five preceding units (1-5) as well as the following three units (7-9). Five of these eight units had been rated as level 4 in structural importance; thus as a generalization of these ideas, it served to summarize several specific and important ideas. Unit 6 is cohesively related to several units. The term "It" is an example of personal text reference, with the referent being the string of clauses preceding it as well as the three following clauses, all of which refer to the lack of appreciation of the sighted and hearing until and if they lose these faculties. "Grateful" is related by lexical cohesion of a reiterative nature in that it is synonymous with "appreciate" and "appreciation" in units 1 and 5. "What" in "what we have" refers to "sight" and "hearing" and involves a demonstrative inference. "Health" involves a collocative reference as a superordinate of "sight" and "hearing" in units 3 and 4, while "ill" can be classified as a superordinate of "deaf" or "blind" in units 1 and 2, and as a synonym for "impairment" in unit 4. This rich textual cohesion contributes to the probability of recall of this generalized macro-proposition.

The most frequently recalled proposition was usually recalled with a low degree of meaning retention, in that only the first part of it was usually paraphrased. Unit 17 said "I who cannot see find hundreds of things to interest me through mere touch." Five of the ten subjects recalling this unit referred only to the fact that the passage was about a person who was blind. Since unit 17 was the only unit in which the author's blindness was referred to explicitly, this recalled idea was matched to unit 17. The rest of the subjects recalling unit 17 stated a paraphrase about the use of the sense of touch or a generalization about the increased awareness of and interest in the environment exhibited by the blind.

The other cognitive operation performed upon the original text in the production of summaries was that of deletion. If a subject had produced a good paraphrase of every text unit, the maximum score would have been 171. The highest summarization score obtained by one of the high cloze scorers was 38, representing approximately 22% of the possible information.

The analysis of cohesive ties within this part of the text could be extended to demonstrate further how closely the propositions in a text are related one to another. As Halliday and Hasan (1976) point out, this type of linguistic analysis does not serve as an interpretation of the text, but helps explain why and how the text means what it does. The analysis of cohesion tells why a text is interpreted a certain way and why it may be ambiguous at certain points. Although readers are not consciously aware of their use of cohesion, it certainly influences their interpretation of the text. Identifying the number and type of cohesive ties provides insight into the probability that certain idea units will be recalled more frequently than others; but like any analysis of discourse features, analysis of texture and cohesion suffers from the inability to account for reader characteristics, such as prior knowledge, and the interaction of these reader characteristics with textual characteristics within the situational context in which the reading occurs. Nevertheless, as noted in the earlier discussion, these ties can help to explain the recall of certain propositions just as they may possibly help to explain why certain words in a cloze passage are, or are not, likely to be filled in correctly. By analyzing how cohesive ties are affected by the deletion of certain words within a cloze passage, and conversely why and how cohesive relations can lead the reader to predict or infer the appropriate word in a cloze passage, it becomes obvious that cloze is affected not only by linguistic cues within the immediate context, but also by longer range contextual constraints, as revealed by the previous analysis of the cohesive ties between unit 5 and the four preceding units.

Before definitive statements can be made regarding the effect of cohesive ties on the probability that certain cloze deletions will be filled in correctly or incorrectly, a complete cohesion analysis of this passage must be carried out and the relationship between number and type of cohesive ties and subjects' cloze responses investigated. If the results demonstrate a positive relationship, then other texts of various discourse types on other topics must also be used in similar analyses. The preliminary analysis in this study indicates that further research in this area is warranted.

ACKNOWLEDGEMENT

No seniority of authorship is intended by the order of authors of this paper.

APPENDIX I

1. _______ Only the deaf appreciate hearing,
2. _______ only the blind realize the manifold blessings that lie in sight.
3. _______ Particularly does this observation apply to those who have lost sight and hearing in adult life.
4. _______ But those who have never suffered impairment of sight or hearing seldom make the fullest use of these blessed faculties.
5. _______ Their eyes and ears take in all sights and sounds hazily, without concentration, and with little appreciation.
6. _______ It is the same old story of not being grateful for what we have until we lost it, of not being conscious of health until we are ill.
7. _______ I have often thought it would be a blessing if each human being were stricken blind and deaf for a few days at some time during his or her early adult life.
8. _______ Darkness would make people more appreciative of sight;
9. _______ silence would teach them the joys of sound.
10. _______ Now and then I have tested my seeing friends to discover what they see.
11. _______ Recently I was visited by a very good friend who had just returned from a long walk in the woods,
12. _______ and I asked her what she had observed.
13. _______ "Nothing in particular," she replied.
14. _______ I might have been incredulous had I not been accustomed to such responses,
15. _______ for long ago I became convinced that the seeing see little.
16. _______ How was it possible, I asked myself, to walk for an hour through the woods and see nothing worthy of note?
17. _______ I who cannot see find hundreds of things to interest me through mere touch.
18. _______ I feel the delicate symmetry of a leaf.
19. _______ I pass my hands lovingly about the smooth skin of a silver birch, or the rough shaggy bark of a pine.
20. _______ In spring I touch the branches of trees hopefully in search of a bud, the first sign of awakening Nature after the winter's sleep.

REFERENCES

Bridge, C. A. *The text-based inferences generated by children in processing written discourse.* Unpublished doctoral dissertation, University of Arizona, 1977.

Bridge, C. A., Tierney, R. J., & Cera, M. J. Inferential operations of children involved in discourse processing. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice,* 27th Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1978.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D. C.: Winston, 1972.

Darnell, D. K. The relation between sentence order and comprehension. *Speech Monographs,* 1963, *30,* 97-100.

Halliday, M. A. K., & Hasan, R. *Cohesion in English.* London: Longman Group Limited, 1976.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic subunits. *Journal of Verbal Learning and Verbal Behavior,* 1970, *9,* 12-20.

Kintsch, W. Comprehension and memory of text. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 5). Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978.

MacGinitie, W. H. *Contextual constraint in English prose.* (Doctoral dissertation, Columbia University, 1960). *Dissertation Abstracts International,* 1960, *21,* 126a-127a.

MacGinitie, W. H. Contextual constraints in English prose paragraphs, *Journal of Psychology,* 1961, *51,* 121-130.

Marshall, N. *The structure of semantic memory for text.* Unpublished doctoral dissertation, Cornell University, 1976.

Ramanauskas, S. *The effect of contextual constraint beyond a sentence on cloze responses of children in special classes for the educable mentally retarded.* Unpublished doctoral dissertation, University of Connecticut, 1971.

Thomas, S. *The effects of prior knowledge upon the contextual constraints in English prose paragraphs.* Unpublished doctoral dissertation, University of Kentucky, 1980.

Tierney, R. J., Bridge, C., & Cera, M. J. The discourse processing operations of children. *Reading Research Quarterly,* 1979, *14* (4), 539-573.

van Dijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just, & P. A. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

PATRICIA BAGGETT
University of Denver

ANDRZEJ EHRENFEUCHT
University of Colorado

CONTENT AND COHESION IN A STORY PRESENTED AS MOVIE VERSUS TEXT

Suppose one wants to compare responses from people who experience a movie or a text that tell the same story, to determine similarities and differences in the resulting memory representations (the information that gets stored in memory) of viewers versus readers (or listeners). Two problems immediately arise. First, what will it mean to say that a movie and text tell the same story? There has been up until now no method to prepare information in two media for experimentation which is the same in content but which can be presented either verbally or pictorially. Forming movie and text stories that convey the same information is necessary in order to make a clear interpretation of similarities and differences in responses. Without such cross media equivalence, response differences may be due to differences in content, and/or differences in the way the two media convey messages. Certainly they may not be unequivocally assigned to media differences alone.

The second problem is to find responses other than, for example, free recall, that test readers (or listeners) and viewers with both verbal and pictorial materials, since the input is verbal in one case and primarily pictorial in the other.

The purpose of this paper is to sketch a series of experiments that solve the two problems above. First, a method will be given to establish a movie and a text that are equivalent in content. Then responses of people experiencing the story in the two media will be described, using both verbal and pictorial materials.

Establishment of Content Equivalence

Content equivalence of two stories will be established empirically using four measures. Further, the *degree* of content equivalence of the stories on each measure will be determined. The stories selected were *The Unicorn in the Garden,* original text by James Thurber (1940), movie directed by William Hurtz (1953). The original Thurber story is 550 words long, or just over three minutes when tape recorded at normal speed. The 16 mm color cartoon is six minutes long.

It would not be very meaningful simply to compare responses from viewers versus readers (or listeners) using the original movie and text as they stand, because the two stories are obviously different in content. For example, the event in the movie of the husband sleepily preparing breakfast is completely omitted from the text. The goal was to match the movie and text better in content. Since it was not feasible to alter the movie, the technique used was to rewrite the text, with certain objectives in mind.

The movie and text were matched as follows:

Identical Openings and Closings

The opening text statement, "Once upon a sunny morning. . .," appears in the first shot of the movie, written on the screen. The final text statement, "Moral: Don't count your boobies until they are hatched," appears in the last shot of the movie, written on the screen.

Identical Dialogue

There are 189 words of dialogue in the original cartoon, and they were put verbatim in the text.

Identical Time Length

The original text was rewritten by a team of three people, using as guides the original Thurber story, a flip book of 180 photos taken at two second intervals from the movie, and 20 free recall protocols from people who wrote them after viewing the movie. The final version of the text was an expanded version of the original and contained almost all of it verbatim. It contained, in addition, descriptions found most frequently in the movie protocols. The important thing for time length equivalence is that, in its final version (1030 words), its tape recorded length equaled the viewing time of the movie, i.e., six minutes. (Copies of the final text are available from the first author.)

Movie and text were matched in one further way which is difficult to quantify. According to Marcus (1977), the visuals of the film capture the style and spirit of the original Thurber drawings, which Thurber created to illustrate his text. In fact, the cartoon merely animates Thurber's illustrations. Both the visuals and the text originated in the mind of the same author, which is unusual when a story is made into a film.

The four measures of content equivalence, and the degree of equivalence on each measure, involve college students' judgments on the movie and the final version of the text. Each of the four will be briefly presented.

Measures Determined by Subjects

Bidirectional Touch Points: Identical Order of Events

A group of subjects, who both saw the movie and heard the text, were given a copy of the text, broken into 350 sentence fragments, and the 180-photo flip book. They selected, for each of 70 starred fragments, the photo that *best corresponded to it in meaning.* The modal photo (the one selected by the most subjects) for each fragment was determined. A second group of subjects, after seeing the movie and hearing the text, selected the sentence fragment that best corresponded in meaning to each modal photo. In 69 of the 70 cases, the modal sentence fragment for the second group was *the original fragment.* In one case, the modal sentence fragment was one removed from the original fragment. Thus a set of 69 perfect *bidirectional touch points,* specific (sentence fragment, photo) pairs that were judged to correspond best in meaning, were established. These touch points also established that the order of events in the two media was identical. (The notion of bidirectional touch points is elaborated on in Baggett, 1977, 1979, where it is introduced).

Representation Ratings

One group of subjects judged how well each of the 350 sentence fragments was represented in the movie, and a second group judged how well each of the 180 photos was represented in the text. On a scale of 0 (not at all represented) to 5 (extremely well represented), the mean ratings of groups one and two were 4.35 and 4.06 respectively, which are not significantly different. They indicate the degree of content equivalence, as assessed by the "how well represented" ratings.

Correlation of Importance Ratings

Subjects heard the text and then judged how important each sentence fragment was to the text. Others saw the movie and judged how important each photo was to the movie. A correlation of the mean importance ratings for the 69 bidirectional touch points was calculated, with $r = .60$. This indicates that sentence fragments and photos that best correspond in meaning have similar importance to the story in their respective media.

Equal Importance of the Same Character in Each of the Media

Subjects who watched the movie rated how important to the movie each of the characters was. Subjects who listened to the text rated each character for importance to the text. The mean ratings for a particular character in the two media were almost identical.

Note that the measures used to determine degree of content equivalence do not require that the materials be in two different media. The measures could be used, for example, on two movies or on two texts. Further, the measures do not require that the material be a story. They ought to work equally well for, e.g., educational materials or instructions.

The content equivalencing done here is not perfect. (See Baggett, Note 1, for a discussion of content differences in movie and final text versions of *The Unicorn in the Garden.*) Still, it is far better in this study than it was in a previous one using movie and text of *The Red Balloon* (Baggett, 1977, 1979). An examination of the data from the equivalence measures above will yield the specific areas in movie and text where the information conveyed is different. The next task is to look at responses, using the content equivalent (to the degree measured here) movie and text as input stimuli.

Experiment: Responses Using Verbal and Pictorial Materials

College student subjects in this experiment watch the movie or hear the text or have no experience with the story. They then do an ordering task, explained below. As will be shown, their errors in ordering can be nicely accounted for by the notions of verbal and pictorial *cohesion.*

METHOD

Subjects

Two hundred forty introductory psychology students served as subjects in partial fulfillment of a course requirement.

Materials

The content equivalent movie and text were the input stimuli. For the task, decks of 23 verbal statements, given in Table 1, and decks of 23 photos from the flip book that are approximate bidirectional touch points with the 23 statements, were constructed. Figure 1 shows 10 of the 23 photos. The photos were in color, 3 in. by 5 in., and mounted on 4 in. by 6 in. index cards. Each photo deck, randomly shuffled, was held together with a rubber band. The statements were typed on 3 in. by 5 in. pieces of paper. Each statement deck, randomly shuffled and held together with a paper clip, was placed in an envelope.

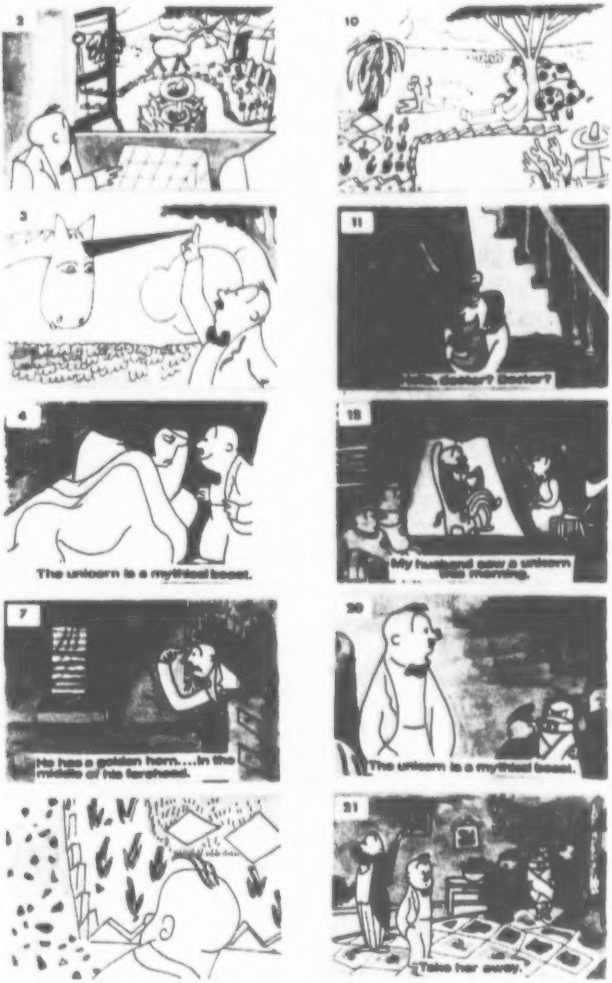

2
10
3
11
doctor? Doctor?
4
The unicorn is a mythical beast.
12
My husband saw a unicorn this morning.
7
He has a golden horn....in the middle of his forehead.
20
The unicorn is a mythical beast.
21
Take her away.

Figure 1. Ten of the 23 photos used in the ordering task, touch points with corresponding statements in Table 1. Numbers in upper left corners indicate the position in the actual story.

TABLE 1
Text Statements which were to be Ordered Correctly
as They Occurred in the Story

1. Husband prepares breakfast.
2. Husband sees unicorn in garden from window.
3. Husband touches unicorn's horn.
4. Wife says the unicorn is a mythical beast.
5. Husband gives unicorn a lily.
6. Wife tells husband he is a booby.
7. Husband says the unicorn has a golden horn in the middle of his forehead.
8. Wife phones police.
9. Husband sees a trail of flower petals.
10. Husband falls asleep.
11. Wife phones psychiatrist.
12. Wife tells psychiatrist and police her husband saw a unicorn.
13. Wife tells psychiatrist and police her husband told her the unicorn ate a lily.
14. Psychiatrist and policemen exchange glances.
15. Wife tells psychiatrist and police her husband told her the unicorn had a golden horn in the middle of its forehead.
16. The psychiatrist signals to the policemen.
17. The policemen seize the wife.
18. The husband sees his wife struggling.
19. The psychiatrist asks the husband, "Did you tell your wife you saw a unicorn this morning?"
20. The husband says the unicorn is a mythical beast.
21. The psychiatrist says to the police, "Take her away."
22. The husband watches the others leave and then grins.
23. Moral: Don't count your boobies until they are hatched.

Procedure

Subjects listen to the text, watched the movie, or had no experience with the story *The Unicorn in the Garden*. Those with story experience were asked, either immediately or after a one week delay, to order the statements or photos correctly in time, as they occurred in the story. Subjects with no story experience were instructed to order the statements or photos so that they "make a good story."

Ten groups of subjects, 24 subjects per group, did the task. These are broken down into five which order statements, and five which order photos of the five groups ordering statments, one has no experience with the story, two hear the text (0- and 7-day delay groups), and two see the movie (0- and 7-day delay groups). The five groups which order photos are formed analogously.

RESULTS AND DISCUSSION

Not one of the 240 subjects did the ordering task correctly! The errors in ordering can be nicely accounted for by the notions of verbal and pictorial *cohesion*.

The claim is that the text statements, ordered correctly as in the actual text (Table 1) have a certain loose cohesion, or clustering together, of statements with repeated words. (See Figure 2.) It can further be shown that orderings of people sorting the statements have a *tighter* cohesion than the original; subjects with no experience have the tightest, followed by 7-day delay subjects second and 0-delay subjects third.

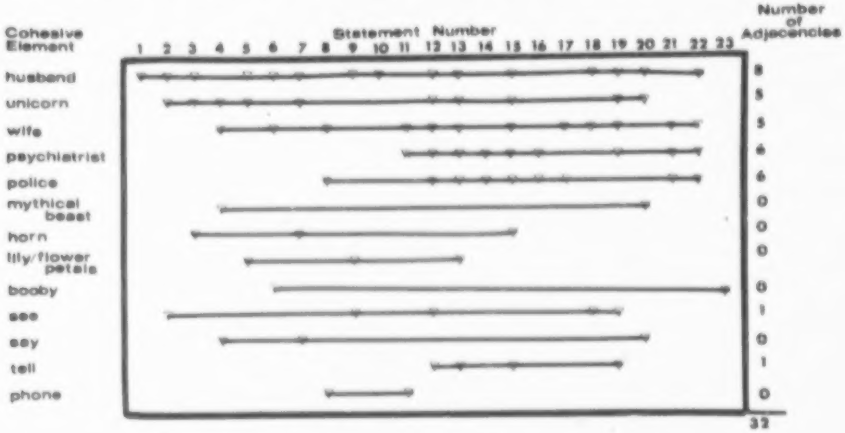


Figure 2. Cohesion graph of text statements in correct order, as they occur in the actual story. Markers indicate that a particular cohesive element occurs in a given statement. The number of times a given element occurs in two adjacent statements is recorded in the far right column under "Number of Adjacencies."

Similarly, the 23 pictorial events ordered correctly as in the movie have a certain somewhat loose cohesion, or clustering together, of photos with similar background locations, characters, props, and dialogue. (See Figure 3.) Further, it can be shown that orderings of individuals sorting photos have a *tighter* cohesion; subjects with no experience have the tightest, followed by 7-day delay subjects second and 0-delay subjects third.

The notion of linguistic cohesion used in this paper comes originally from Halliday and Hasan (1976) and was quantified by Ehrenfeucht (1977), who introduced the idea of a linguistic *cohesion graph*. We extend it here to *pictorial cohesion* as well.

To derive the linguistic cohesion graph for the 23 sentences, consider each sentence as a *unit*, and each single word (noun or verb) that occurs in at least two sentences as a potential *cohesive element*. (A noun and its pronoun referent are also included as cohesive elements. See Figure 2.) For each element, a marker is placed in the unit where it occurs. To measure cohesion, we decided that if a given element occurs in two adjacent units, it is maximally close. So we count the *number of adjacencies* for each element, and take the sum over all elements. For the statements there are 13 cohesive elements and, in their correct story order, there are 32 total adjacencies.

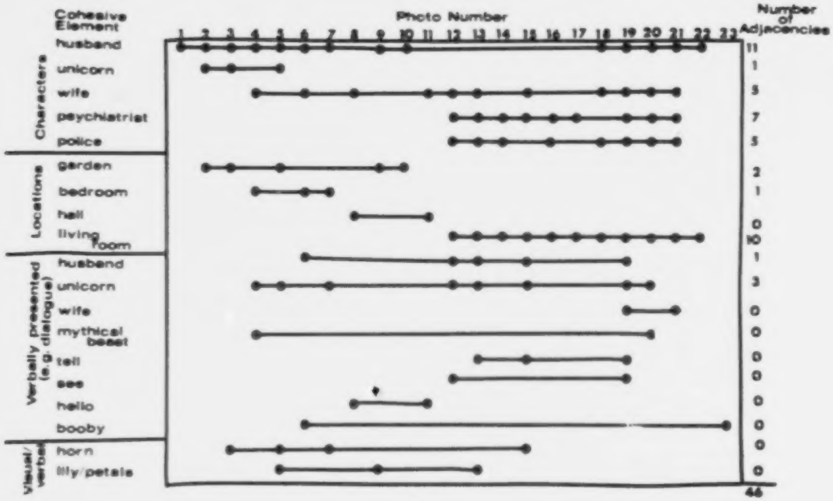


Figure 3. Cohesion graph of movie photos in correct order, as they occur in the actual story. Markers indicate that a paticular cohesive element occurs in a given photo. The number of times a given element occurs in two adjacent photos is recorded in the far right column under "Number of Adjacencies."

Each individual's order can similarly be scored for the number of adjacencies. This is done by forming a cohesion graph with the person's ordering of statement numbers across the top, rather than the correct order as in Figure 2, and counting the number of adjacencies for that person's order. (The scoring was done by computer.) The mean number of adjacencies for each of the five groups who ordered statements is presented in Table 2. All five of these scores are significantly greater than the 32 adjacencies in the correct order.

TABLE 2

Mean Number of Adjacent Cohesive Elements
for Subjects Ordering Text Statements

Input Medium	Delay Zero	 7-day
Text	38.7	40.25
Movie	39.1	40.5
No story experience	42.2	

Note. There are 32 adjacent cohesive elements in the correct order. Twenty-four subjects were in each group.

A pictorial cohesion graph can be derived similarly. Each of the 23 photos is considered a unit, and the 19 cohesive elements selected, which occur in the surface of the photos, are as shown in Figure 3. They fall into the

TABLE 3
Mean Number of Adjacent Cohesive Elements
for Subjects Ordering Movie Photos

Input Medium	**Delay**	
	Zero	7-day
Movie	52.9	54.2
Text	52.4	55.1
No story experience	55.1	

Note: There are 46 adjacent cohesive elements in the correct order. Twenty-four subjects were in each group.

categories of characters, locations, verbally presented words, and elements presented both verbally and visually. As before, the number of adjacencies can be calculated for the original order (46, as shown in Figure 3), and for each individual's order. The mean number of adjacencies for each of the five groups ordering photos is given in Table 3. All five of these scores are significantly greater than the 46 adjacencies in the correct order. On examining the five categories of cohesive elements in the photos individually, it appears that *locations* (which are *not available* in the text statements) are more significant clusterers than any of the others.

It can be concluded that the groups with no story experience substantially place together or cluster cohesive elements in their orderings. The groups with story experience seem caught between a memory for the correct order, which is strong but imperfect at zero delay and weakens over seven days, and a tendency to be influenced by the cohesive elements present in the statements or photos. This tendency increases considerably from 0- to 7-day delay. The adjacency scores of the 7-day groups and those with no story experience are almost identical.

The medium of input (movie versus text) for the groups who experience the story does not have much influence on the adjacency scores. Doing the ordering task in the modality identical to the one at input gives no advantage, at either delay.

Two possible mechanisms could explain the finding highest cohesion by people with no story experience and a shift toward tighter cohesion over a delay by people with story experience. First, clustering may be purely a processing phenomenon that happens at the time of task: subjects cluster when they see similar elements. Or second, associations in memory may be made by an awareness of the existence of similar elements, e.g., as postulated by Kintsch's theory (Kintsch and vanDijk, 1978; Kintsch and Vipond, 1978).

Both of these mechanisms lead to the conclusion of the tighter clustering over a delay. For explanation (1), subjects try to cluster as much as they can. But the actual recollection of the events in the story prevents clustering. The more subjects forget, the more they cluster.

Another factor that prevents clustering is of course world knowledge that, e.g., a goes before b, or a moral goes at the end of the story. Such constraints account for the fact that subjects' orderings do not necessarily have the tightest possible cohesion.

For explanation (2), memory is reorganized over time, and connections between the units which are cohesive story close, while connections between units that are not cohesive fade away.

The experiment reported here does not decide between the two mechanisms, but predictions from the two can be checked experimentally, by varying cohesive elements independently in the story input and in the task stimuli. Then clustering by the cues in the task versus by previous memory cues can be compared. Depending on which of the two types of cues clusters, the mechanisms which is operating can be determined. If cues in the *task stimuli* are clustered, then mechanism (1) is operating. If not, and cues in *memory* are clustered, then mechanism (2) is operating.

The result found here is important because how subjects order the photos and statements might depend on any one or several of very many possible factors. We have determined according to which principles they cluster, and our cohesion analysis predicts what errors subjects will make, namely, in the direction of a greater number of adjacent cohesive elements.

Summarizing, the main findings using the response measure of ordering photos or statements are that, as memory for order decreases, cohesion increases; and presenting information in different media, provided it is content equivalent in the two media, does not make a difference. This does not prove that the conceptual information stored in memory from content equivalent movie and text is in the same format. But it does mean that, if the information is not in the same format, communication between different formats proceeds without difficulty.

FINAL REMARKS

Hopefully the methods for establishing content equivalence outlined here will be useful in future picture-text research. They offer a new way to assess content differences in pictures and text. Further, the cohesion analysis is a new way to look at text and picture order and how cohesion changes when order changes. It is hoped also that it will be a new and useful tool for data analysis.

ACKNOWLEDGMENTS

This work was supported by a Faculty Development Award from Florida International University and NIMH postdoctoral fellowship number 1 F32 MH07588-02 to the first author. Sincere thanks go to Renee Scheiber and Firmin Bernabe for helping to rewrite the original Thurber story and for aiding in data collection. Also, much appreciation is expressed to Pam Hoge, the artist who created the black and white drawings for Figure 1.

REFERENCES

Baggett, P. *The formation and recall of structurally equivalent stories in movie and text.* Unpublished doctoral dissertation, University of Colorado, 1977.

Baggett, P. Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior,* 1979, *18,* 333-356.

Ehrenfeucht, A. *How to tell a text from a nontext.* Talk given to Semiotic Circle, University of Colorado, February 2, 1977.

Halliday, M., & Hasan, R. *Cohesion in English.* London: Longman Press, 1976.

Kintsch, W. & vanDijk, T. Toward a model of text comprehension and production. *Psychological Review*, 1978, *85*, 363-394.

Kintsch, W. & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In Lars-Goran Nilsson (Ed.), *Proceedings of the conference on memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Marcus, F. *Short story/short film*. Englewood Cliffs, N.J.: Prentice Hall, 1977.

Thurber, J. The unicorn in the garden. In F. Marcus (Ed.), *Short story/short film*. Englewood Cliffs, N.J.: Prentice Hall, 1972.

RANDOLPH K. CIRILO
University of Texas at Austin

SOME EFFECTS OF STORY STRUCTURE DURING COMPREHENSION

Recently, there has been a great upsurge in work dealing with story comprehension. A major aspect of this research has followed from the development of theoretical systems which attempt to describe the underlying structure of stories (e.g. Rumelhart, 1975; Mandler & Johnson, 1977; Thorndyke, 1977). In order to validate these notions about structure for the casual reader, researchers have typically compared subjects' memory representations for comprehended stories with structural descriptions for the same stories, looking for corresponding properties. For example, a differentiation of so-called high-level and low-level information has been taken as evidence that, during comprehension, a structured representation is built for a story. This representation goes beyond simple intersentential links like anaphora.

However, because such research relies on memory data, the time course of the processes that would yield such representations is unclear. The present research attempts to circumvent this problem by tapping comprehension more directly. The purpose of this work is to discover if the structural properties of stories also influence the immediate processing of their component sentences.

EXPERIMENT I

The first experiment focuses on the hierarchical nature of story structures (see Cirilo & Foss, 1980, for a more complete treatment). Many experiments have shown that high-level information is remembered better than low-level information (e.g. Thorndyke, 1977). It was reasoned that high-level sentences might also receive longer reading times than low-level sentences. This would follow if subjects tend to allot more attention to high-level propositions during processing or if high-level propositions tend to be more difficult to integrate with the previous context. Concurrently, pilot work indicated that sentences occurring early in a story might receive longer reading times than those occurring late, so this factor was also included in the experiment.

METHOD

Materials

Quadruples of stories were constructed. Across each member of the quadruple, a particular sentence was repeated varying both its hierarchical level, high or low (in accordance with Thorndyke's, 1977, story grammar), and its serial position, early or late (the first or second half of the story). High-level sentences were always in the first or second levels of their particular structures, while low-level sentences were separated by at least one level in their story structure from that occupied by the corresponding high-level sentence. Three such quadruples were written, ranging in length from 27 to 38 lines. The reading times for the repeated sentences across the four conditions were the data of interest.

Procedure

Twenty-four college students read the quadruples along with four practice stories. The stories were presented one line at a time on a computer-controlled display scope. Subjects controlled the rate of presentation by pressing a button which eliminated the line currently on the screen and replaced it with a new line directly below it until the bottom of the screen was reached at which point presentation shifted back to the top. The time between button presses served as a measure of reading time. The order in which stories were presented varied across subjects, although the members of each quadruple were always separated by two members of the other quadruples. Finally, after all the stories had been read, recall protocols were collected for three of them. These were scored for the presence or absence of the gist of the critical sentence in the story read.

RESULTS AND DISCUSSION

The results are shown in Table 1. As expected, significant effects of both hierarchical level, $F(1,23)=12.15$, $p<.005$, and serial position, $F(1,23)=6.12$, $p<.025$, were obtained. High-level sentences took longer to read than low-level sentences and early-occurring sentences took longer than late-occurring sentences. Moreover, the recall data supported the structural positions assigned to the critical sentences. They were recalled 66% of the time when they were high-level as compared to 26% of the time when they were low-level.

TABLE I

Mean Reading Times (msec) for Critical Sentences

	Serial Position	
Hierarchical Position	**Early**	**Late**
High	1966	1785
Low	1837	1602

As stated earlier, the effect of hierarchical level lends itself to more than one explanation. Readers may be capable of determining the important propositions in a story on the basis of cues in the text as well as general expectations about the ways stories are put together. Consequently, more time might be devoted to such propositions in order either to integrate them well with other information in memory or to help make sure they receive long-term storage. Alternatively, high-level sentences could simply be more difficult to integrate with the previous context. They could be less predictable in that they tend to introduce new material rather than expand upon material already presented. Unlike the first account, this type of reading time difference would not be related to the memory difference. Research is currently underway to determine the nature of the processing which results in this reading time effect.

As for the effect of serial position, it too lends itself to multiple interpretations. The speedup could reflect increasing contextual constraints or simply a greater density of information early in stories as opposed to late. Again, research will be needed to differentiate such possibilities.

EXPERIMENT II

Despite the above ambiguities, the effect of hierarchical level on reading time argues that a story's structure influences comprehension as well as memory. The second experiment deals with a second property of story structures that might be expected to influence reading time. Specifically, the plot of a story is said to consist of a sequence of one or more episodes. A number of recent experiments (e.g. Glenn, 1978; Mandler, 1978) have shown that abstractly defined episodes act as constituents in memory, determining how information will be organized and remembered. It was asked, then, if this structure might also be evidenced in reading time data.

A result obtained by Thorndyke (1977) suggested that this might be the case. He found that subjects' memory for a story was better when it was read following a story that duplicated its structure than following a story with a different structure. One explanation offered for this effect was that readers were able to use the already learned structure of the first story in order to aid their understanding of the second story. In other words, less time was required to build the structure for the second story when it duplicated a previously built structure.

It seemed that a similar argument might hold for the individual episodes within a single story. If, during processing, sequences of text are being chunked into episodes, these substructures within the larger structure of the story as a whole might in some sense be available to the comprehender. Consequently, they might be used to facilitate the processing of similar substructures which follow them. In short, just as the processing of a second story was facilitated because it shared the structure of a previous story, so might the processing of an episode in a multi-episode story be facilitated because it shared the structure of the one before it. If it is, in fact, read more quickly in such a case, there is evidence for the creation of episodes during processing itself.

METHOD

Materials

Twelve two-episode stories were written, the episodes ranging in length from eight to sixteen lines. Moreover, there were two versions of each story. In one, the two episodes had the same structure. That is, for every event or description in the first episode, there was a corresponding event or description in the second. Similar intersentential connections and syntactic forms were also used. In the other version, there was no such parallelism in structure. These manipulations were performed only with respect to the contents of the first episode while leaving everything else unchanged across versions.

Episodes are defined here as temporally conjoined sequences of events with potentially recognizable beginnings and ends. They were not always simple goal-structures as in story-grammatical treatments. However, all were amenable to the structural analysis referred to in describing how the two versions of each story were written.

Procedure

Twenty-two college students read 12 of the 24 possible stories using the same procedure as in Experiment I. In half the stories they read, the two episodes in each story had a common structure, while in the other half the two structures were different. This was counterbalanced across subjects by using the two versions of each story. That is, the six stories with a common episodic structure and the six with a different structure for one subject were read by another subject using the opposite versions. The reading times for the lines of the second episode were the data of interest. Recall protocols were also obtained for three of the twelve stories read.

RESULTS AND DISCUSSION

The results are shown in Table 2. The times are divided into quintiles of the second episode, so it could be determined if any effects found varied with position in an episode. Because episodes varied in line length (see above), the size of a quintile also varied across stories. The average size of a quintile was 2.7 lines. The quintile was arbitrarily chosen, since the episodes could not be more conveniently parsed into segments.

Significant effects of both structure (identical versus different), $F(1,21)=9.13$, $p<.01$, and quintile, $F(4,84)=19.67$, $p<.001$, were obtained. As expected, the average reading time for an episode was shorter when it repeated the structure of the episode prior to it. Also, reading times appeared to decrease as subjects got farther into an episode. However, the number of words per line also decreased with quintile, so this may have been the responsible factor. It might also be noted that although the effect of structure seemed to decrease across quintiles, the test for an interaction was not significant. As for the recall data, somewhat more lines were remembered from the second episode of common structure stories than from those with differing episodic structures. This is what one would expect from Thorndyke's results. The difference was small (47% to 39%) but reliable, $t(11)=2.22$, $p<.05$, when analyzed by story. However, not enough recall data were collected for us to put much confidence in this result.

TABLE II

Mean Reading Times (msec) for Lines in Each Quintile of the Second Episode

	Quintile				
Structure of Episodes	**1**	**2**	**3**	**4**	**5**
Different	1895	1744	1795	1576	1615
Shared	1727	1663	1656	1498	1610

In short, the results support the hypothesis that subjects are creating an episodic structure for stories during comprehension. Moreover, this process must be fairly immediate, since the facilitation due to shared episodic structure is manifest in even the first quintile of the second episode. However, the nature of the facilitation is far less clear. Is the subject in some sense directly comparing the structure in the second episode with that in the first; or might the result reflect the priming of those processes which were utilized in the comprehension of the first episode? Whatever the case, the results provide evidence that the constituent structure of a story has an influence on its immediate processing.

To summarize, Experiments I and II have demonstrated a direct influence of structural properties, hierarchical level and episodic structure, during the course of reading. Thus, structural variables can be studied during comprehension and not just at retrieval as in most previous studies. Research of this sort, then, may prove to be a fruitful area for the future.

REFERENCES

Cirilo, R. K., & Foss, D. J. Text structure and reading time for sentences. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 96-109.

Glenn, C. G. The role of episodic structure and story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 29-247.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1978, *1*, 14-35.

Mandler, J. M., & Johnson, N. J. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Rumelhart, D. E. Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.), *Representation and understanding*. New York: Academic Press, 1975.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

JANE L. DAVIDSON
DEWAYNE TRIPLETT
Northern Illinois University

THEMATIC CONVERGENCE IN RECONSTRUCTED PASSAGES IN MAPPING ACTIVITIES

The current interest in studying the process of comprehension has led investigators in numerous directions, such as, studies involving the reader's interaction with text and analysis of discourse. A major implication from studies is the need for instructional strategies which maximize the reader's comprehension of text. A reader who can retell a passage may or may not be able to comprehend the passage. The value of having students actively involved in the reading and thinking process has been urged by reading authorities (Henderson, 1977; Huey, 1908; Russell, 1964; Stauffer, 1980). Opportunities for discussion before and after reading encourage different kinds of thinking. "When discussions vary in the story dimensions considered, when different kinds of thinking processes are encouraged, and when they become true occasions for dialogue, then students have a stimulating thinking experience" (Hall, 1978).

A procedure called the Group Mapping Activity (Davidson & Bayliss, 1978) provides for dynamic interaction through discussion after reading. Mapping was defined by Hanf (1971) as a means of graphically structuring information and was used as a substitute for note taking or outlining. Elements of ideas and concepts provided the basic structure of the map and were identified by skimming the passage. After reading the selection, relationships between ideas and concepts were drawn on the map from memory.

Davidson and Bayliss (1979) recognized that Hanf's procedure involved not only recall of information but other higher thought processes as well. They modified the procedure for use as a group activity after reading and called it the Group Mapping Activity. Students within the group mapped their perceptions of the passage and shared their reasons for drawing in particular elements. The sharing atmosphere was intense; questions and responses represented higher levels of thinking. Questions asked by group members were analyzed by means of the Ohio Scales (Wolfe, Huck, & King, 1967) and were determined to be chiefly interpretive and evaluative type questions. Students interacted with other group members and gained insights about the various processes used by others in mapping their perceptions of the passage read. A puzzling phenomenon which appeared frequently in some of the maps which were drawn indicated that some students shifted the protagonists. There appeared to be some common patterns in other maps.

The purpose of this study was to analyze and classify maps and determine whether the protagonist shift occurred more frequently in any one pattern of classification. The four classifications of maps are as follows:

1. Interactive — the map shows a major emphasis on interactions between and relationships among characters in the passage.
2. Identification—the map shows a major emphasis on one character in the passage.

3. Linear — the map shows a major emphasis on the sequence of events in the passage.

4. Miscellaneous — the map shows fragments or irrelevant elements pertaining to the passage.

Four types of internal elements in the classifications of maps are:

1. Character relationships — elements indicate relationships between and among characters in the passage.

2. Symbolic awareness — elements indicate symbolism or abstract references.

3. Sequence of events — elements identify the sequence of events in the passage.

4. Ancillary characters — elements identify ancillary characters in the passage.

In order to minimize difficulties in literary interpretations for subjects in this study, the Directed Reading-Thinking Activity (Stauffer,1969) was used in presenting the passage to be read. Squires (1964) found that readers experienced difficulties interpreting literary passages when they failed to grasp the implications of details or made incorrect inferences or failed to defer final judgment until they related to details presented later in a story. He identified the need for readers to obtain all information before making final assessments. The DR-TA is a group problem-solving procedure in which the teacher facilitates the readers' thinking by regulating the amount of material to be read and by asking questions which stimulate their examination of available clues and declaration of purposes for reading. In literature, purposes for reading are stated as predictions or expectations of the development of the story and final outcome. Purposes differ for individual readers, depending on their background of experiences, the experiences of the group, and the content of the material. Stauffer stated, "Purposes for reading represent the key element in versatility" (1969, p. 24). He also stated that where readers have "become involved in the dynamics of a purpose-setting session, the self-commitment on an intellectual as well as an emotional level has tremendous motivating force" (1969, p. 25). During the DR-TA, the readers skim, scan, and study as they read in order to satisfy their purposes. They check on individual and group purposes as they read and either keep their original predictions or redefine them. Thereby, concepts in the passage may be further developed. Moreover, the group discussion in which predictions were confirmed, discarded, or redefined would insure that subjects participating in the study would have shared understanding of the passage.

The final step of the DR-TA usually involves discussion, further reading, additional study, and writing in order to increase powers of observation, reflection by abstraction, judgment, and reasoning. For the purposes of this study, the Group Mapping Activity was used as the follow-up activity or final step to the DR-TA.

METHOD

Subjects who participated in the study were 59 undergraduate and graduate students in education enrolled in reading and language arts courses taught by the authors. The short story, "All the Years of Her Life," by

Morley Callaghan was presented to all students by means of the Directed Reading-Thinking Activity. After students had completed the reading of the passage, they were instructed to retell the passage in writing. The Group Mapping Activity was explained to the subjects and they were instructed to map their perceptions of the passage. The written retellings and maps were collected by the examiners.

The subjects' written retellings were analyzed for recall of information from the passage, which was parsed according to Stein and Glenn's (1979) story grammar. It must be noted that: 1) subjects in the study were reading the material at their independent levels of reading achievement; 2) the DR-TA procedure was used to present the material to be read; and, 3) the task of written retelling involved immediate recall of information. (Ten written retellings were arbitrarily selected at random to examine whether the retellings showed the subjects were comprehending at their independent levels of reading achievement. Results of this preliminary investigation showed that the ten subjects had no difficulty in recalling information, and the remaining 49 subjects' written retellings were then examined.) Analysis of the 59 subjects' written retellings showed that all the subjects had no difficulty recalling information. No subject shifted the protagonist on the written retelling of the passage.

SUMMARY OF RESULTS

An analysis of the maps yielded the four classifications: Interactive, Identification, Linear, and Miscellaneous. A total of 59 maps were evaluated. Fifty-eight maps were classified in one of the three observed patterns; one map was not decipherable; it was labeled Miscellaneous. Interactive patterns were mapped most frequently. In this category, 23 (.39) maps were classified. Identification patterns were observed next most often with 19 (.32) maps classified in this group. The Linear patterns were observed in 16 (.27) of the maps.

A review of the percent of map classifications showing the protagonist or protagonist shift indicated that a total of 50 (.84) maps identified the protagonist, and eight (.14) maps revealed a protagonist shift. One (.02) map was classified as Miscellaneous.

Of the 23 maps classified Interactive, a total of 22 (.37) identified the protagonist; one map identified a protagonist shift. In the Identification classification, 12 maps (.20) identified the protagonist and seven (.12) identified a protagonist shift. The 16 maps (.27) classified as Linear identified the protagonist.

An analysis of internal elements in maps classified as Interactive, Identification, and Linear was conducted. Internal elements consisted of four basic types: Character Relationships, Symbolic Awareness, Sequence of Events, and Ancillary Characteristics. Analyses of the 58 maps yielded 79 internal elements that were categorized among the three map classifications and the four types of elements. The total of 79 internal elements were distributed among the three map classifications as follows: Interactive, 38; Identification, 25; and Linear, 16. The internal elements were distributed among the four types of elements as follows:

1. Character Relationships: A total of 29 internal elements involved relationships among or between characters. Maps classified as Interactive

contained 22 internal elements which showed character relationships. The Identification maps contained seven internal elements which showed character relationships. There were no internal elements of this type in the maps classified as Linear.

2. Symbolic Awareness: A total of 23 internal elements involved symbolic awareness. Maps classified as Interactive contained ten internal elements which showed symbolic awareness. The Identification maps contained 12 internal elements which showed symbolic awareness. There was one internal element of this basic type in the maps classified as Linear.

3. Sequence of Events: A total of 18 internal elements involved the sequence of events. Maps classified as Interactive contained 18 internal elements which showed the sequence of events. The Identification maps contained two internal elements which showed the sequence of events. There were 13 internal elements of this type in the maps classified as Linear.

4. Ancillary Characters: A total of nine internal elements involved ancillary characters. Maps classified as Interactive contained three internal elements which showed ancillary characters. The Identification maps contained four internal elements related to ancillary characters. There were two internal elements of this type in the maps classified as Linear.

DISCUSSION

Results of the study suggest that maps can be analyzed and classified as Interactive, Identification, or Linear. The protagonist shift appeared most frequently in the maps classified as identification. Readers who drew Identification maps perceived the protagonist (or shifted-protagonist) as being most important. The internal elements in these maps represented strong emphasis on symbolic references (i.e. hardships, emotions) as well as emphasis on the protagonists' (or shifted-protagonists') relationships with other major characters in the passage.

Readers who shifted the protagonist when they drew their maps had not shifted the protagonist in their written retelling of the passage. As one subject commented, "I know what the story is about, but I don't care. The mother in the story is really what's important to me." It's as if the subject was saying to the examiners, "If you ask me for recall, I can give you what you want, but my perceptions will remain my own."

Maps classified as Interactive showed that subjects were primarily concerned with interactions and relationships among the characters in the passage. These maps contained internal elements which represented strong emphasis on relationships and interrelationships identified or suggested in the passage. Although symbolic elements were emphasized, they were of less importance.

The results of the examination of maps classified as Interactive or Identification suggested that these readers reacted emotionally as well as intellectually to information in the story.

Readers' maps classified as Linear showed little evidence of any emotional involvement, intellectual awareness of relationships among characters, or symbolism in the passage. Their maps provided evidence of primary concern with the sequence of events only.

Maps provide a means of gaining additional insights about the ways in which readers organize information which can be accounted for by schema theory. "Schema theory provides a way of integrating our understanding of text with our understanding of the world in general" (Adams & Collins, 1979, p. 21). The readers in this study demonstrated their abilities to store information in memory and immediately recall it; moreover, the results of analysis of map protocols show that readers' organization of information is selective.

This study is considered exploratory. Further investigation is needed to determine whether individuals are consistent in the types of maps they draw after reading other types of materials. Additional study also needs to be directed toward gaining greater understanding of readers who shift the protagonist as they organize information. One investigation might explore the long term recall of the reader who shifts the protagonist in order to determine if the reader's personal response to another character is of sufficient strength to blot out accurate recall of the passage. Spiro (1977) discussed conditions for inaccurate recall and suggested that "the likelihood of inaccuracy increases with the degree of interaction with preexisting cognitive structures. The more one relates discourse to prior knowledge, the greater the assimilation into those structures with the resultant loss of particular identity of the discourse" (p. 139).

The Group Mapping Activity provides a means by which individuals' personal responses to reading can be examined in order to better understand the impact of readers' interaction with text. Adams and Collins point out that, "At the highest level, one has successfully read a passage if one understands it both as it was intended by the author and in terms of its impact on oneself" (Adams and Collins, 1979, p. 1).

REFERENCES

Adams, M. J., & Collins, A. A schema-theoretic view of reading. In R. Freedle (Ed.), *New directions in discourse processing* (Vol. 2). Norwood, NJ: Ablex Publishing Company, 1979.

Callaghan, M. All the years of her life. *The New Yorker*, June 8, 1935.

Davidson, J. L. The DR-TA: A reading thinking strategy for all levels. In R. T. Vacca & J. A. Meagher (Eds.), *Reading as a language-experience.* Storrs, CT: The Reading-Study Center, The University of Connecticut, 1978.

Davidson, J. L., & Bayliss, V. A. Mapping. *The Journal of Language Experience*, 1979, *1*, 7-15.

Hall, M., Ribovich, J. K., & Ramig, C. J. *Reading and the elementary school child* (2nd ed.). New York: D. Van Nostrand Company, 1979.

Hammond, W. D. The personalization of comprehension. In A. Guenther (Ed.), *IRA Language-Experience Special Interest Group Newsletter*, 10-13, 16.

Hanf, M. B. Mapping: A technique for translating reading into thinking. *The Journal of Reading*, 1971, *4*, 225-230, 270.

Henderson, E. H. The role of skills in teaching reading. *Theory Into Practice*, 1977, *16* (5), 348-356.

Huey, E. *The psychology and pedagogy of reading.* Cambridge, MA: M. I. T. Press, 1908 (1968).

Meyer, B. J. F. Organization in prose and memory: Research with application to reading comprehension. In P. D. Pearson (Ed.), *Reading: Theory, research and practice.* Twenty-sixth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1977.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, NJ: Erlbaum Publishers, 1977.

Squires, J. R. *The responses of adolescents while reading four short stories*. Champaign, IL: National Council of Teachers of English, 1964.

Stauffer, R. G. *Directing reading maturity as a cognitive process*. New York: Harper and Row, 1969.

Stauffer, R. G. *The language experience approach to the teaching of reading*. New York: Harper and Row, 1980.

Stauffer, R. G. *Teaching reading as a thinking process*. New York: Harper and Row, 1969.

Stein, N. L. & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multidisciplinary perspectives in discourse comprehension*. Norwood, NJ: Ablex, 1978.

Wolfe, W., Huck, C.S., & King, M. L. *Critical reading ability of elementary school children*. U. S. Department of Health, Education, and Welfare, Office of Education, Project No. 5-1040, Contract No. OE-4-10-187. Columbus: The Ohio State University, 1967.

111

THOMAS H. ESTES
WAYNE L. SHEBILSKE
University of Virginia

COMPREHENSION: OF WHAT THE READER SEES OF WHAT THE AUTHOR SAYS

This paper derives from an NIE/NSF jointly funded project designed to develop some basic understandings concerning the relationships between the conceptual structure of textbooks and the quality and processes of comprehension. Many students have difficulty learning from the textbooks of content courses, though sometimes the students are otherwise seemingly good readers. At least two plausible explanations might be offered for this. For one, it is possible that most of the materials in which students have learned to read possess a narrative structure, while those in which they are now asked to read to learn are characterized by an expository structure which makes different demands on the reader. After several years of being taught to read in anthologies of stories, students may have tacitly internalized "grammars" for stories, strategies of comprehension not dissimilar in principle from those which researchers such as Rumelhart (1977) have explicated in the form of narrative rules. The expository and quasi-technical writing in textbooks, particularly at upper grade levels, may not easily yield to such rules.

Another explanation is that the world knowledge demands placed on students by their content area reading are prohibitively great. Narration, with its major purpose of effect rather than information, would seem to make relatively fewer such demands. If so, this might argue for its choice as the style for teaching very young readers, but the question remains, how are students to understand the more difficult expository form of textbooks whose major purpose is to inform? Our working hypothesis is that the answer to this question will be found in an explication of what Walter Kintsch (1979) has called an "interactionist" view of reading. Our hope is to contribute to this explication.

The base of research from which one begins to investigate the process of comprehension is quite solid. Two particular features characterize research in comprehension and retention of prose in this decade and distinguish it from previous work. The last ten years have witnessed a change from surface structure analysis of text to the analysis of semantic structure, concomitant with a switch to open-ended or minimally cued recalls in preference to answers to specific questions as a measure of comprehension (cf. Marshall & Glock, 1978). These shifts reflect a widespread view that there are (1) identifiable structures of text, generally specifiable in propositional terms, and (2) structures of thinking, hypothesized about in relation to cognitive structure (Ausubel, 1968) and schema theory (Anderson, 1976), which both have an effect on comprehension and, indeed, may together play a determining role with respect to comprehension. One of the major assumptions to which this leads is that what the reader *sees* in the text plays as heavily in determining comprehension as does what the author *says* in the text. In philosophical terms, this idea is congruent with the phenomenological perspective on reading advanced most recently and persuasively by Wolfgang Iser in his book, *The Act of Reading:* ". . .the study of a literary work should concern not only the actual text but also, and in equal measure, the actions involved in responding to that text" (Iser, 1978, pp. 20-21. For a shorter and easily apprehensible discussion of Iser's ideas, see Iser, 1974, or Iser, 1971.)

Much notable recent research has rested on ideas similar to these (Thorndyke, 1977; Rumelhart, 1977; Kintsch & van Dijk, 1976; Frederickson, 1975). In general, however, most researchers have sought to infer the contribution of the reader by looking for differences between the structure of text and the structure of recall. One of our major attempts to extend this line of inquiry will be to develop tools to quantify what the reader *sees* in the text. We seek to advance the operation of text analysis by giving it the power to account as directly as possible for the contribution of the reader to the structure of text. Though we are utilizing formal analyses of text structure, the emphasis of our research is on the structure of text as readers perceive it.

Our work has begun with textbooks in science. We have specified in propositional terms the formal structure of excerpts from three biology textbooks, one at the seventh grade level, one at the tenth grade level, and one at the college level, all concerned with the classification of species. We have employed a dependency analysis proposed by James Desse (1980) which yields a propositionally-based hierarchical explication of text. McCabe (in press) has used this analysis in describing children's narratives, and her paper provides a particularly nice discussion of the procedure for anyone wanting to see the method described in detail in the context of one of its uses. The units of analysis are syntactically defined propositions. The theoretical foundation of the procedure is Chomsky's (1965) grammatical theory involving a syntactic base and a set of transformations that modify that base. The transformations of interest in Deese's analysis are those that interrelate propositions by establishing equivalences, coordinations, or subordinations between propositions. The result of the analysis is the de-transformation of these transformations into a hierarchical outline of the discourse. That is, the analysis displays discourse as a hierarchy. McCabe (in press) describes the procedure: "(1) isolate the most dominant proposition; (2) determine the propositions, if any, that are contrastive within that (dominant) proposition, break them out, and subordinate them to the dominant propositions so that they are parallel and coordinate with each other; (3) repeat this process with every segment of the discourse until all contrastive propositions have been de-transformed."

From this analysis, two important though not unexpected findings have emerged in our research. First, the passages we are using differ in their depth. The seventh grade passage has no proposition deeper than the sixth level in the hierarchical structure. In contrast, the tenth grade passage has propositions as deep as the twelfth level. Second, the texts differ in the extent to which they demand inferences on the part of the reader, the extent to which information is presented ambiguously, and the extent to which the organization of the presentation follows an inductive or a deductive path.

For example, the subsection of the tenth grade text entitled "Meaning of species" begins with a particular example, the "many intermediate mongrels between such different breeds as Great Danes and greyhounds". The first paragraph ends, "Therefore, all dogs are grouped into a single species." The full section ends with a one sentence definition of species as its summary. The general problem of speciation is buried in the text and in the hierarchy of propositions which describe it. Other sections are even more extremely inductive, and the reader is often left to infer very major ideas. These show up as parenthetical propositions in the formal analysis. By contrast, however, the corresponding section of the college text, entitled "What is a species?" begins with a definition. Similarly, other sections tend to begin immediately with general principles, and thus the organization of the text is such that the topical statement is formed of very high order propositions. Exactly how this may affect what the reader sees in the text is one of our major questions. In technical terms, this relates to the grammars assumed by different expository structures.

Our formal analyses are also directed at the problem of world knowledge presumed by text. Though lower level texts typically have less depth of structure than higher level texts, this is not invariably so. For example, the excerpts on the classification of species are atypical because the excerpt from the college text has a more shallow structure than the excerpt from the tenth grade text. Yet our intuitions suggest that the college excerpt is actually more complex to understand. The beginnings of an answer to this enigma are found in consideration of the part of the specific body of knowledge presented by each text. The college text relates the notion of species to the theory of evolution; the tenth grade text does not, the authors having chosen to treat evolution as a separate topic. Hence, the information presented in the college text must be related by the reader to a larger conceptual structure which is mostly assumed by the author. The difference between the two texts is no simple matter of sentence complexity, word frequency, or even propositional content. It is rather, at least in part, a matter of what we might call "textual presumption", the relationship between the conceptual structure of each text and the structure of an entire body of knowledge about biological classification. It is part of our goal to adopt the use of formal analyses in a manner which allows us to capture this important conceptual feature of texts.

The complement to these formal measures of text properties, and the dimension of our research allowing an approach to text as a product of readers' perceptions, thus far has taken the form of four interrelated behavioral measures. Initial data were drawn on college subjects; similar measures are currently being taken on students at seventh and tenth grade levels. The data reported here were derived from college students' reading of our tenth grade passage. College students have served, in effect, as expert readers and it is from data drawn on their reading of the tenth grade text that we expect to derive hypotheses concerning other students' reading of materials appropriate to their grade placement. These are our four measures: (1) We have asked students to mark "idea units" in the text and we have analyzed these markings to determine where the texts divide into such units. (2) We have asked a group of students to rate these idea units as important or unimportant with respect to the author's main points. (3) We have asked a group of students to read the passage and to record immediately their recall of what they read. (4) We have asked a group of college students to rate the importance of the propositions ensuing from our dependency analysis with respect to the author's central message. Interrelationships among these data suggest interesting connections between the manner in which students perceive and act on text structure and the structure of their comprehension.

One immediately interesting finding is that formal and behavioral measures may differ in the picture of text they afford. Any formal analysis must be specifiably and predictably rule governed, by definition, and in that degree it must be rigid. Behaviors of readers, though, are governed by more flexible rules, rules which are tacit and unspecifiable in principle. This can result, for example, in some high level propositions being given low importance ratings and in some low level propositions being given high importance ratings. Most obviously, this occurs when good readers perceive the importance of subsidiary propositions which are main ideas.

Our procedure for determining the idea units in a text is interestingly simple but complexly justified mathematically (Rotondo, 1979). Basically, the problem we faced was in knowing how many idea units a passage contained and where the breaks between units came. To answer the first question, we relied on the arithmetic mean of idea unit markings which subjects used. That is, our best guess of how many ideas the text contained was the average number of ideas marked. To answer the second question, we calculated the percentage of subjects agreeing on each possible idea unit break, ranked the possible breaks on that basis, and then merely counted off the number of ideas our averaging had told us exist in the passage.

A hypothetical example might serve to clarify these procedures. Assume a passage of narration on a certain topic such as a chapter from a textbook. Assume further that a group of 100 readers is asked to read the passage and to place a slash mark at each point where they think one idea ends and another begins. Different readers will have different numbers of idea units thus marked, naturally. But the best estimate of how many ideas units probably exist in the passage is the average number of units the 100 readers chose. Say the average number of ideas marked in this passage is 20. The question remains, which are the "actual" 20 which the passage contains? The answer is, those the 100 readers had the highest agreement on. Which possible place for a slash mark was most freqently chosen? Next most frequently? And so on, until the 20 most frequently chosen idea units are determined. The hypothetical passage is now divided into idea units based on readers' perceptions. The procedure is elegantly simple and lets us describe the idea unit constituents of any text based on a group of readers' perceptions. This represents a major purpose of our research which we will continue to refine in practice.

Our first major research question has concerned the relationship between the importance rating of individual idea units and the likelihood of their inclusion in the recalls of students who read the passage. As one might expect, we find a substantial and significant correlation between importance and recall of an idea ($r = 0.39$, $p \leqslant .001$). This accounts, however, for only 16% of the variance, leaving us to conjecture about the other 84%. We are particularly interested in two kinds of deviations: 1) students failing to report what they know to be important and 2) students tending to report what they know to be unimportant. These particular anomalies suggest something to us about the use of examples, allowing us to go beyond the simple notion that concrete examples facilitate memory (cf. Gibson & Levin, 1975). First, we are now hypothesizing that concrete examples do not facilitate understanding when readers misperceive the connection between examples and the abstract ideas they are chosen to illustrate. For instance, all of our subjects failed to recall, or perhaps even to understand, the important idea that geographical isolation can maintain distinctness between species. The illustrative example was of polar bears and brown bears which don't interbreed in the wild because they never see each other. (They will do so in zoos, if allowed.) Our readers tended to remember details about the living habits of the two kinds of bears, but failed to include in their recalls the main point of that detail. As a consequence we are now working with teachers to refine strategies by which to teach students to find connections between interesting detail and more important generalizations.

As a second hypothesis arising out of these data, we are predicting that some examples are so compelling and interesting in their own right that they tend to be recalled even when they illustrate relatively trivial points. We are now testing the hypothesis that readers will recall main ideas better when trivial ones are not highlighted by their particularly vivid illustration.

The results of our research project will have implications both for what authors should say and for helping readers see text more clearly. For example, authors should take care not to bury main points deep in the hierarchy of their ideas or leave them to be inferred. Examples and main ideas should be clearly connected. Compelling examples should be reserved for important ideas. But readers have a big responsibility as well to adequately interpret what they read and to learn from text despite its ambiguity and lack of clarity. One of our major goals is to help teachers find ways to improve skills of identifying and remembering the main points of text.

REFERENCES

Anderson, R.C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Frameworks for comprehending discourse* (Tech. Rep. 12). Champaign, Ill.: Center for the Study of Reading, University of Illinois, July 1976. (Eric Document Reproduction Service No. ED 134 935)

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart, and Winston, 1968.

Chomsky, N. *Aspects of the theory of syntax.* Cambridge, Mass: M.I.T. Press, 1965.

Deese, J. *Specification of dependency relations in text.* Unpublished manuscript, University of Virginia, 1980.

Frederiksen, C. H. Representing logical and semantic structures of knowledge acquired from discourse. *Cognitive Psychology,* 1975, *1,* 371-458.

Gibson, E. J., & Levin, H. *The psychology of reading.* Cambridge, Mass: M.I.T. Press, 1975.

Iser, W. Indeterminancy and the reader's response in prose fiction. In G. H. Miller (Ed.), *Aspects of narrative.* New York: Columbia University Press, 1971.

Iser, W. *The act of reading: A theory of aesthetic response.* Baltimore: Johns Hopkins University Press, 1978.

Iser, W. The reading process: A phenomenological approach. In R. Cohen (Ed.), *New directions in literary history.* Baltimore: Johns Hopkins University, 1974.

Kintsch, W. *On comprehension.* Paper presented at the meeting of the American Educational Research Association, San Francisco, April 1979.

Kintsch, W., & van Dijk, T. A. *Recalling and summarizing stories.* Unpublished manuscript, University of Colorado, 1976.

Marshall, N., & Glock, M. D. Comprehension of connected discourse: A study into the relationships between the structure of text and information recalled. *Reading Research Quarterly,* 1978-1979, *14,* 10-55.

McCabe, A. Organization of propositions in narratives. In A. McCabe & C. Menig-Peterson, (Eds.), *Three psycholinguistic analyses of children's narratives.* New York: Plenum Press, in press.

Rotondo, J. *Clustering analysis of subjective partitions of text.* Unpublished manuscript, University of Virginia, 1979.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processess in reading: Perception and comprehension.* Hillsdale, N. J.: Lawrence Erlbaum, 1977.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology,* 1977, *9,* 77-110.

V. FRANK STONE
Western Illinois University

THE EFFECT OF INTERSENTENTIAL RELATIONS ON THE COMPREHENSION OF CONNECTED DISCOURSE

Comprehension lies at the center of the reading process, and in recent years the major focus of inquiry for reading research has been directed toward the various components of comprehension. A prominent topic has been to ascertain what variables of text affect its comprehensibility and thus the comprehension of text by the reader. The present investigation was directed toward the semantic and syntactic relations between sentences within a passage of connected discourse. The influence of intersentential relations has been shown to affect the flow of processing information in sentence pairs (Haviland & Clark, 1974; Clark & Haviland, 1977) and in five-sentence paragraphs (Carpenter & Just, 1977).

INTERSENTENTIAL RELATIONS

Haviland and Clark (1974) and Clark and Haviland (1977) constructed sentence pairs to observe the effect of intersentential relations on processing time. They proposed that readers incorporate information contained in previous sentences in comprehending successive sentences. Their results suggest that when a writer does not provide the reader with a direct antecedent for new information presented in a text, the reader is forced to spend more time relating the two sentences together — constructing a cohesive tie. Thus, it appears that readers utilize information from previous sentences in the comprehension of successive sentences.

Recent theories on comprehension have been directed toward the verb and its associated cases (Fillmore, 1968; Kintsch, 1974; Norman & Rumelhart, 1975; Shank, 1973). To analyze the function of semantic and anaphoric relations between sentences, Carpenter and Just (1977) examined the differences in latency time for processing target sentences following sentences containing verbs with direct inference case entailments and sentences with indirect inference case entailments. The verbs were selected so that each entailed verb had a corresponding nonentailed verb (e.g.), if someone is assassinated, then a killer is implied, but if someone is said to have died, that does not entail or imply a killer. Carpenter and Just hypothesized that sentences containing a verb and a direct entailment (assassinated/the killer) would take less time to integrate than those sentences containing a verb and an indirect entailment (died/the killer). Their results strongly suggest that readers carry information from previous sentences in comprehending the successive sentences of five-sentence paragraphs.

The semantic representation of a verb provides a cohesive element that binds sentences together to form a text. These intersentential relations appear to play an important role in the comprehension of sentence pairs and five-sentence paragraphs. There is no indication from prior research, however, as to what readers will do when confronted with an ambiguity (indirect inference) within a longer passage of connected discourse. The goal of the present investigation was to extend the results of the empirical studies of Carpenter and Just to typical reading material.

METHOD

Subjects

A group of six female and six male undergraduate and graduate students at the University of Virginia served as subjects for this study.

Apparatus

The eye movements of readers were recorded by an EDL/Biometrics Reading Eye II machine. This device records on a printout the pattern of eye movements of reading while they are engaged in the reading act.

Materials

The materials used were two passages from ***The Reader's Digest*** condensed readers. One passage, "Remember the Alamo," contained 1,971 words and the other passage, "Appomattox: Epic Surrender," contained 1,516 words. The passages were typed on 5x9 cards in elite type with 35±10 single-spaced lines. The cards were then placed on the EDL/EYE II machine to be read. The "Alamo" passage was on ten cards and the "Appomattox" passage was on seven cards.

The independent variable was condition of the entailed verb. The two conditions were direct inference case entailment (DI) and indirect inference case entailment (II). The "Alamo" passage contained 14 entailed verbs and the "Appomattox" passage contained 6 entailed verbs. The condition, direct or indirect, of the 20 entailed verbs was randomly ordered within the two passages. There were a total of 10 entailed verbs in the direct inference case entailment condition and 10 entailed verbs in the indirect inference case entailment condition. The first ordering of conditions within the two passages became material set A. To provide a means to compare the effect of entailed verbs on reading time, it was necessary to compare the same target sentence in each condition. To accomplish this, a second order of conditions was constructed — material set B. Material set B consisted of the same two passages as material set A with the same number of entailed verbs. However, the condition of the entailed verbs was opposite of those in material set A. Thus, the two material sets could have had the following order of conditions:

Material Set A

Entailed verb	Condition
rushed	D
spotted	I
went	I
argued	D
chased	D
injured	I
painted	D
stabbed	D
sat	I
blasted	D
requested	I
galloped	D
death	I
communicated	I
swept	D
returned	I
walked	I
retreat	D
insisted	D
began	I

Material Set B

Entailed verb	Condition
went	I
encountered	D
rode	D
discussed	I
went	I
shot	D
depicted	I
died	I
slept	D
replied	I
insisted	D
came	I
assassination	D
wrote	D
made	I
rode	D
climbed	D
move	I
asked	I
opened fire	D

Two sample paragraphs follow, one from material set A and one from material set B, showing the two conditions as they were read by subjects:

Material Set A

A few survivors made their last stand in small rooms of the old church. Crockett met his end in one of these rooms. Around his body were 17 dead Mexicans. Propped on his elbow in his cot, Bowie fired his pistols. While trying to raise his knife to defend himself, Bowie *died*. The bayonets plunged into him from all directions.

Material Set B

A few survivors made their last stand in small rooms of the old church. Crockett met his end in one of these rooms. Around his body were 17 dead Mexicans. Propped on his elbow in his cot, Bowie fired his pistols. While trying to raise his knife to defend himself, Bowie was *stabbed*. The bayonets plunged into him from all directions.

The last sentence in each of the above paragraphs was the target sentence for which reading times were calculated. The entailed verbs in the examples were in the penultimate sentences.

Analysis

The dependent measure was the reading time for the target sentence following the entailed verb. The raw data from the printout, in milleseconds, was converted to the dependent measure, words per minute.

A repeated measure-mixed design was utilized to analyze the converted data. The within subjects factor was the average reading time of the target sentences in the direct and indirect case entailment conditions. This factor provided an opportunity to compare how the same subjects read sentences in both direct and indirect conditions. However, each subject did not read each target sentence in both conditions.

The between subjects factor was material set A versus material set B. The two material sets were counterbalanced so that all target sentences were read equally often after direct and indirect case entailment verbs and all combinations of passage order.

The results of the analysis revealed that there was a distinct difference between the reading times of those sentences preceded by a direct inference case entailment verb. The sentences preceded by a direct inference case entailment verb had a mean reading time of 360.80 words per minute while those sentences preceded by an indirect inference case entailment verb had a mean reading time of 295.20 words per minute. The mean difference of 65.60 words per minute between the two conditions was significant, F (1,10) = 23.15, $p < .001$. The mean reading times for the two material sets were 324.00 words per minute for material set A and 332.08 words per minute for material set B. The two sets of materials were not significantly different, F (1,1) = .04, $p > .05$. There was no interaction between the two independent factors, i.e., direct vs. indirect inference case entailment verbs and the material sets A and B, F (1,1) = 1.25, $p > .05$. The mean reading rate for material set A was 349.83 words per minute in the direct condition and 298.83 words per minute in the indirect condition. Likewise, the mean reading rate for material set B was 372.50 words per minute in the direct condition and 291.66 words per minute in the indirect condition.

IMPLICATIONS

The most important finding of this study was that readers, in processing information from natural text (i.e., language in use), carry information across sentences using the preceding information in an effort to comprehend successive sentences. This finding is consistent with the major hypothesis of the study and prior research using sentence pairs and five-sentence paragraphs (Carpenter & Just, 1977; Clark & Haviland, 1977; Haviland & Clark, 1974). The importance of textual cohesion in comprehension, (i.e., verbs and associated case entailments), can now be viewed within an environment of larger passages of discourse. Like the previous investigation suggested, even within larger texts, readers can comprehend a sentence more rapidly when it is cohesively connected rather than inferentially connected by means of argument repetition to a preceding sentence (Kintsch & Vipond, 1978).

The results of the present inquiry provide further evidence that linguistic structure assists readers in their integration of information from a text base into their message base. That is, whenever a bridge (Clark, 1976) is required by an incoherent text, its construction will reduce the reader's processing capacity during comprehension, hence, affecting reading time and comprehensibility of the text. In addition, the results add supportive evidence to the hypothesis made by Miller and Kintsch (1980) that places within a text that are not cohesive will cause additional processing difficulties for readers. Lastly, prior to this investigation, it was not known how readers would react when encountering incoherence within a typical passage of connected discourse.

This study of intersentential relations within connected discourse has extended the paradigm incorporated by Haviland and Clark and Carpenter and Just to a rich and more realistic context of reading. This opens the possibility of the study of specific additional contextual variables within this framework. Knowing that the nature of intersentential relations does affect reading times of sentences in context, we can now proceed to manipulate certain contextual and extra-textual variables in order to observe other types of cohesive ties in finer detail. These variables might include, for example, purpose and reading strategies, set, prior knowledge, inferences, coherence, and structure of text.

REFERENCES

Carpenter, P. A., & Just, M. A. Reading comprehension as eyes see it. In P. A. Carpenter & M. A. Just (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Clark, H. H., & Haviland, S. E. Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension: Discourse process, advance in research and theory*. Norwood, N.J.: Ablex, 1977.

Clark, H. H. Inferences in comprehension. In D. LaBerge & S. J. Samuels (Eds.), *Perception and comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Fillmore, C. J. The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart & Winston, 1968.

Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 512-521.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Ed.), *Memory: Processes and problems*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Miller, J. R., & Kintsch, W. *Readability and recall of short prose passages: A theoretical analysis*. Paper presented at the American Educational Research Association Conference, Boston, Mass., April 1980.

Norman, D. A., & Rumelhart, D. E. *Exploration in cognition*. San Francisco: W. H. Freeman, 1975.

Schank, R. C. Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: W. H. Freeman, 1973.

DR. CAROL A. HODGES
University of Illinois

READING COMPREHENSION INSTRUCTION IN THE FOURTH GRADE: THREE DATA-GATHERING METHODS

Mastery of reading is awarded a place of importance as a goal of education by the public and professional educators as well as by the federal government (Shami & Hershkowitz, 1974). While considerable research has focused on the acquisition of various reading skills considered to be important to the act of reading (Huey, 1908; Anderson & Dearborn, 1952; Gibson & Levin, 1975), most of that knowledge has concerned decoding while much less is known about the processes involved in and the teaching of reading comprehension. It has only been within the last decade or so that research has begun to provide theoretical concepts which may give us a more definitive understanding about the nature of reading comprehension.

From the point of view of teacher education, it is imperative that this new research provide the field with resources to improve comprehension instruction. If theory and basic research are to make a difference in classroom instruction, questions regarding their appropriateness, relevance, and usefulness to the classroom teachers' way of understanding reading comprehnsion need to be answered.

Attempts to change curricula have often run up against long established prejudices and procedures built into schooling. Results which did not take into consideration the assumptions or constraints of school or which ran counter to prevailing opinion have often been dismissed as invalid or ignored entirely (Petrie, 1968; Goodlad & Klein, 1974). It appears that we have been saying that we know what schools need without knowing what schools already have; and that we know what to put into schools without knowing what needs to be replaced. This study begins a process of discovering what schools and teachers are doing in the area of reading comprehension.

PURPOSE AND RESEARCH QUESTIONS

The purpose of this study was to document, through data gathered from surveys, interviews, and obervation, the approaches that teachers in fourth grade in public elementary schools were using in order to help children understand what they read. The study focused on the *actual* behaviors, methods and materials of the teacher in the classroom rather than on what reading methods, texts, or basal series manuals say *should* be done. This was done with the firm belief that one way to facilitate the acceptance of new instructional procedures based upon research is to provide researchers with information regarding current classroom practices. Such information can provide researchers with knowledge about the variables that teachers believe in and the terminology through which they express ideas about those variables. In turn, researchers can use this information to illustrate to teachers how new instructional practices based on research are compatible with what the teacher believes to be important or to test new practices against existing ones and present the results to teachers in ways that illustrate their appropriateness, relevancy, and usefulness.

The first major research question of the study was: What approaches (behaviors, methods, and materials) do fourth grade teachers in public elementary schools use to help children understand what they are reading? There were four related sub-questions:

a. How much time is spent in formal reading instruction each day in the classes studied?
b. What kinds of reading activities take place in that time span?
c. What kinds of materials are used? How are they used?
d. How are the classes organized for reading?

The second major question was: What are the teachers' concepts of reading comprehension? There were three related sub-questions:

a. How do teachers define comprehension?
b. What are the teachers' sources of information regarding comprehension.
c. How do teachers assess comprehension in their classrooms?

METHODOLOGY

Sample

The sample population for this study consisted of 44 fourth grade teachers in public schools in East Central Illinois who taught reading in self-contained classrooms or departmentalized settings. This sample resulted in a variety of fourth grade teachers, seven male and thirty-seven females, some who had less than a year teaching experience and others who had more than 20 years experience. Teachers from both rural and urban school districts were included.

Procedures

After the sample had been identified, questionnaires were sent to the 44 teachers. Eight of the teachers who responded to the survey were randomly selected to be interviewed. Another eight were selected for a three-day observation of their classroom instruction in reading. Four members of this latter group were also interviewed.

Instruments and Processes

Instruments were designed for each step of the data gathering. The survey and interview questionnaires focused on each of the sub-questions related to the major research questions.

The observation schedule was constructed to obtain information about the teachers' actual instructional programs in reading and a content area. It allowed recording of data regarding the time taken to instruct, and the goals, methods, and materials under consideration during that time. The interactions which were recorded were then classified according to pre-specified categories. Thus both the raw data and summary data were available for analysis.

Analysis

The data were first analyzed and presented in a quantitative manner utilizing data gathered by each of the sources. In this way data gathered on each sub-question could be compared and contrasted across the survey, interview, and observation data-gathering methods. Through the use of tables, figures, and verbal summaries, information was gained regarding the kinds of reading skills teachers regard as most important to teach and the methods and materials they use to teach those skills.

The second method of presenting data consisted of four qualitative descriptions of how teachers present instruction in reading. These short case studies or mini-tales were compiled from a synthesis of the data gathered from four teachers who participated in all three data-gathering methods.

RESULTS

For the sake of brevity the findings from three methods are summarized and combined.

I. What approaches do fourth grade teachers in public elementary schools use to help children understand what they are reading?

a. How much time is spent in formal reading instruction each day?

The 44 teachers spent from 40-120 minutes of their school day in formal reading instruction. The average time spent was 60 minutes which represented 18 percent of the average school day. Teacher self-reports were usually in terms of an average time, which often did not present a clear story of the kinds of factors which impinge on the actual program. Interviews and observations illustrated that many teachers have their reading period scheduled by extraneous factors such as the need to coordinate teaching with another teacher, the requirements of other subject matter, or the fact that the organizational scheme utilized by the majority — that is, grouping in conjunction with a basal series — usually demanded at least 20 minutes of teacher time per group. The observations also revealed that as much as 19 percent of the time of some reading lessons was spent in non-instructional activities such as disciplining, passing out paper, or changing seats.

b. What kinds of reading activities take place in that time span?

c. What kinds of materials are used? How are they used?

d. How are the classes organized for reading?

The teaching of literal comprehension was reported by teachers and confirmed by interviews and observation as the most important instructional activity in fourth grade. This was followed by the teaching of word recognition and word meaning skills. Other instructional activities such as inferential comprehension, study skills and rate of reading were ranked much lower and more or less equally behind the top three.

Three teachers who shared a large class of 75 students organized their classroom in a flexible manner around the use of tradebooks and skill groups. The other 41 teachers used a basal series as the developmental core of their program. Of those teachers who used basals, four taught reading by using one group, 13 organized two groups, 11 teachers had three reading groups, and nine teachers used four groups. In addition there were two teachers each who used five and six reading groups. By and large these groups proved to be inflexible after the first three weeks of school. Twenty-seven teachers reported that they used only one basal series in their program while 14 used two or more of them as enrichment or remedial resources. Twelve of the teachers who reported using basal series commented that they did nothing to modify the basal teachers' guide but used it as it was prescribed. Five teachers reported that they used *no* materials other than basals; but this was not the case in the majority of the classrooms. Other materials which were used included tradebooks, reading "labs," skillbuilders, various commercial and teacher-made games, and dittos. These materials were abundant in some classrooms and practically non-existent in others.

The use of oral reading followed by oral questions was the most widely used method of utilizing material. Silent reading of the material followed by oral questions was second. Basal workbook assignments as specified by the teachers' manual was the third most often used method of teaching. By combining the information gathered thus far, one can begin to see a pattern emerging which illustrates that the majority of the teachers viewed the oral or silent reading of basal series followed by oral questions as the best way to reach their number one teaching goal — literal comprehension. Very little *direct verbal teaching* of any comprehension skill, whether thought of as literal or inferential was reported or observed.

II. What are the teachers' concepts of reading comprehension?

a. How do teachers define comprehension?

When the investigation focused exclusively on comprehension, teachers agreed that one of the most important goals was to help children understand what they read. However, they tended to define comprehension in terms of the skills associated with it, not in terms of its processes or other variables deemed by researchers to be important to comprehension. Variables such as preparation, modeling, oral language facility, or reinforcement were seldom mentioned. As the content of their comprehension instruction was discussed, it became clear that the majority of teachers provided instruction with regard to those skills which are generally classified as literal comprehension by researchers while excluding those classified as inferential comprehension.

b. What are the teachers' sources of information regarding comprehension?

The basal teachers' guide ranked first as the source of comprehension information. Most teachers stated that their reading methods courses had done little to prepare them to teach comprehension. It appears that whatever is suggested by the basal teachers' manual, as the instruction for the day, is followed by most teachers. Little diagnostic teaching takes place.

c. How do teachers assess comprehension in their classroom?

The same technique used to teach comprehension was also widely used to assess it; that is, teachers asked oral questions after the children read the material. In addition, many teachers reported using the basal worksheets as evaluative devices. One comment frequently heard was that although children could do well on the worksheets, they could not accomplish the same task in a "real reading situation". It appears that much instruction and assessment is done out of context of reading situations; and therefore there is little transfer of learning to "real reading".

CONCLUSIONS AND IMPLICATIONS

The conclusions which can be drawn from such data provide strong implications for reading researchers and teacher educators who are interested in improving teachers' knowledge about the reading comprehension process and in improving the manner in which reading comprehension can be taught in our schools. The data from this study illustrate that although the majority of teachers who participated believed that comprehension was important to reading, they were uncertain about what comprehension and comprehension instruction entailed. Furthermore, the data show that the commercial materials upon which the teachers depend for the majority of their instruction (predominately basal readers) tend to approach comprehension from a point of view of a series of skills to be mastered separately. However, the various approaches to studying comprehension which are being researched strongly suggest that comprehension is a multi-dimensional process involving the holistic cognitive processing of language.

Implications for Teacher Educators

One of the most useful features of having derived a definition of reading comprehension instruction from instructional theory is that the resulting data can link up fairly directly with teacher education programs. One theme which ran through the results was the fact that many teachers simply have a lack of knowledge about the complex nature of reading comprehension and its instruction. While we may already discuss such factors as diagnostic teaching, study skills, etc., in our methods courses, the data from this study were collected using categories that enable us to see immediately which aspects of instructional theory are being utilized and which need more emphasis. Although the current research-based theory is still incomplete, there is some knowledge about reading comprehension that can make a difference in how a teacher instructs by providing a framework for making day-to-day instructional decisions.

Multiple factors influencing comprehension. Because teachers appeared vague about factors other than skills instruction which influenced comprehension, reading methods instruction might well begin by providing teachers with a simple model which explicates the important variables in teaching comprehension. Once the theory behind these factors has been introduced, practical ideas for using them in instruction can be presented. Furthermore, strategies for promoting comprehension such as concept stretching, previewing, feedback, and modeling can be demonstrated.

Diagnostic teaching. It is apparent that the majority of teachers studied depended upon basal reading series to tell them how and when to teach a topic. The immediate needs of their students appeared to be secondary. While the basal reader may act as a backbone for instruction, teachers must realize that they are the ones who are best able to diagnose what their students need at any particular time. Because the basal reader appears to dominate classroom instruction, teachers need to receive instruction on how to use the basal in a diagnostic manner.

Direct verbal instruction. In addition to supplying information to teachers regarding learning *when* to teach a topic, teacher educators need to supply more information about *how* to teach a topic. The importance of direct verbal instruction should not be underestimated. Most exercises in basal texts and workbooks *test* students on their knowledge of comprehension skills rather than *teach* them to use the skill in everyday reading. The general ability of the student to think does not imply that he will be aware of how or when to do it. Without direct instruction there is no reason to believe that a student will be able to infer or read critically.

Question asking. Teachers relied heavily upon questions to direct and to evaluate the level of the students' understanding of the material. Teacher educators must realize that effective questioning is not an innate talent, but rather it is a skill that can be developed with practice. Instruction in a reading methods course should contain information regarding both the types of questions and the sequences which can serve different purposes and stimulate different levels of thinking.

Implications for Reading Researchers

In addition to helping teachers become more knowledgeable about current research findings, reading researchers must consider ways in which they can disseminate information about future findings.

Efficient use of time. Almost any new instructional practice will make demands on the teacher that the basal approach does not make. A common theme that ran through the results was the complaint by teachers that they did not have enough time to teach in the ways they would like to teach. If new instructional practices are to be accepted, it is imperative that researchers use such studies as the present one to find out how much time teachers have available and how they are currently using their time. Then the researcher can be prepared to illustrate to the teacher how the new instructional practice will be useful and not more time consuming.

Basal reader collaboration. As long as teachers rely so heavily on the basal reader as their source of instruction, it appears that reading researchers should investigate the possibility that new instructional practices can be disseminated most effectively through the basal reader. While it may not be easy to convince basal publishing companies to move from "tried and true" methods, an attempt must be made. Many veteran teachers will never take another college reading course. One way to continue the teachers' education, however, is for researchers and basal publishers to join hands in accepting that responsibility.

Bridging the language gap between researchers and practitioners. To gain acceptance of new instructional procedures by teachers, it seems imperative that the language describing the procedures be that which teachers can readily understand. The researcher who can show how his new procedure, whether it be from a psycholinguistic, cognitive psychology, or other viewpoint, relates to the teacher's idea of comprehension will be the researcher whose new procedures will be more readily accepted. The researcher should begin with the teacher's definition of comprehension as "understanding what is read," and add to that definition details that utilize instructional theory concepts with which the teacher is familiar. If the researcher casts the new procedure in such terms, the teacher is more likely to see how it will fit into her programs.

Investigations such as this one, which document for researchers and teacher educators some of the language and realities of classroom life, may provide the beginnings of the bridge linking the fruits of research with the acceptance of practical instructional concepts.

REFERENCES

Anderson, I. H., & Dearborn, W. F. *The psychology of teaching reading.* New York: Ronald Press Co., 1952.

Gibson, E., & Levin, H. *The psychology of reading.* Cambridge, Mass.: The MIT Press, 1976.

Goodlad, J. I., Klein, M. F., & Associates. *Looking behind the classroom door.* Washington, Ohio: Charles A. Jones Publishing Co., 1974.

Huey, E. B. *The psychology and pedagogy of reading.* Cambridge, Mass.: MIT press, 1968. (Originally published, New York: MacMillan Co., 1908)

Petrie, H. *What has learning theory failed to teach us how to learn? Philosophy of Education, 1968.* Proceedings of the Twenty-fourth Annual Meeting of the Philosophy of Education Society, 163-170.

Shami, M. A., & Hershkowitz, M. Relative importance of mastery of reading skills as a goal of public education. *Journal of Reading Behavior,* 1974, *6,* 98-99.

CINDY STEINBERG
BERTRAM BRUCE
Bolt Beranek and Newman Inc.

HIGHER LEVEL FEATURES IN CHILDREN'S STORIES: RHETORICAL STRUCTURE AND CONFLICT

INTRODUCTION

Traditional surveys of children's literature have examined features such as text structure and topic, but have failed to take into account rhetorical elements such as author-reader distance, commentary, point of view, and insight into characters' minds. Similarly, they have glossed over aspects of character to character interaction, such as responses to interpersonal conflict. These "higher level features" of stories may be what make stories interesting to read. They are also principal contributors to story complexity, and hence, to difficulty for beginning readers. On both counts, interest and complexity, it is important to come to a better understanding of these features.

To concretize our discussion, we first present two examples showing the importance of higher level text features. Second, we sketch a theory of higher level story features. Then, we briefly describe how we are applying our analysis to a selection of children's stories. Finally, we discuss some implications of this work.

EXAMPLES

To illustrate a higher level text feature which we have studied in children's stories, let's consider the following modified fairy tale:

> Once upon a time there lived a little girl who always wore a red cloak with a big red hood. One day her mother asked her to take a cake to her ailing grandmother. Being fond of her grandmother, she put on her cloak and joyfully started out on the errand. The little girl took a path through the woods gathering nuts and flowers along the way. By and by, she reached her grannie's cottage. When she knocked at the door the old woman welcomed her in. They both sat down and had some cake together. After they were finished, the little girl said goodbye and went home.

Someone who heard this story might well say, "And what happened?" Clearly, something is missing from the story, even though the words, the sentence structures, the characters and the topic are similar to those in "Little Red Riding Hood." But who would remember Little Red Riding Hood today if she had never encountered the wolf? It seems that conflict is an essential ingredient in this story, and perhaps for stories in general. As the novelist John Le Carre (Barber, 1977) says, "The cat sat on the mat, is not a story. The cat sat on the dog's mat is a story."

Conflict makes us feel that a story is indeed a story. Furthermore, the types of conflict we see in a text affect our enjoyment and understanding of it. Thus, although conflict has traditionally been viewed as a topic within the domain of literary analysis and criticism, it also has a proper role in the study of how children learn to read. If we are to determine what features of texts contribute to comprehensibility, memorability, and reading pleasure we need to understand better what conflict is and how it operates in stories.

In a similar way, we have been led to consider other features of texts that may have important influences on comprehension and motivation. Among these, we have focused particular attention on features which derive from the relationship between an author and a reader. The popular children's story "The Tale of Benjamin Bunny" by Beatrix Potter exemplifies important elements in that relationship.

The text begins with a description of Benjamin Bunny's trip to the woods to visit his aunt, Mrs. Rabbit and cousins—Flopsy, Mopsy, Cotton-tail and Peter. Setting information, character description, and character interaction are conveyed almost entirely in third person narrative. However, notice the narrative twist in the following excerpted paragraph:

> Old Mrs. Rabbit was a widow, she earned her living by knitting rabbit-wool mittens and muffatees (I once bought a pair at a bazaar). She also sold herbs, and rosemary tea, and rabbit-tobacco (which is what we call lavender).

Who is "I" and "we" in this example? How can the reader understand these comments and their relationship to the rest of the story? If the reader were to view the text as encompassing another level of social interaction—an author to reader level—beyond that of the obvious character interactions, he or she would go far in constructing a coherent and plausible explanation for these phrases. The "I" then becomes the author, Beatrix Potter, who speaks directly to the reader. She invites the reader to believe in the rabbit world which she has created—a world where rabbits sell muffatees at bazaars and people buy them. Potter even goes further in her attempts to communicate with the reader through the use of "we" referring to the reader, Potter herself, and presumably other humans or non-rabbits.

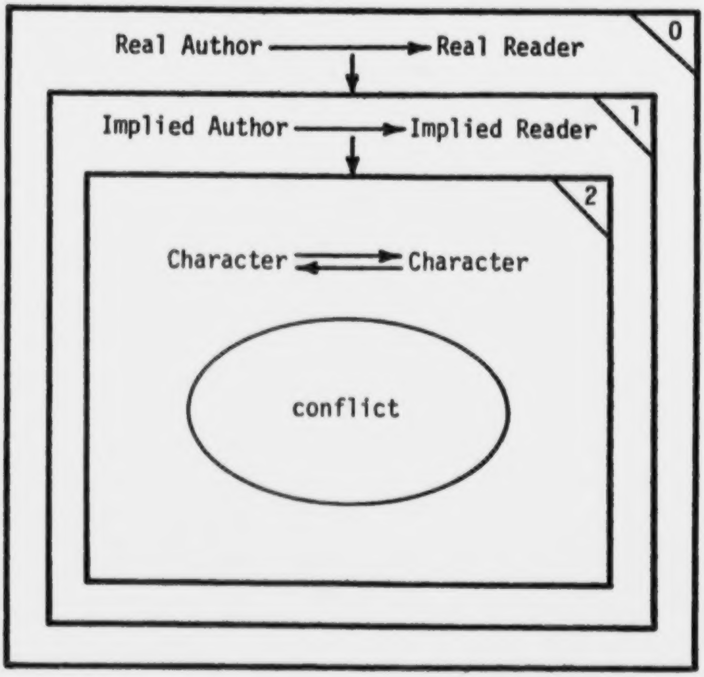
Beatrix Potter makes extensive use of direct author to reader communication. This style may contribute to the success of her stories. Her technique serves to draw the reader into the story, diminishing the distance created by the physical text. Direct address makes the reader feel as if gaining his or her attention is important to the author. It is an acknowledgement on the part of the author that the readers' views and concerns are relevant to the story.

Author to reader text communication serves important functions in addition to the enhancement of reader involvement. Author commentary often adds meaning to a text through the elaboration or explication of story events. Key story ingredients such as foreshadowing, suspense, surprise, and satire are frequently conveyed through the utilization of some form of the author-reader relationship. We believe that recognition of the central role this aspect of text plays in stimulating reader interest and enhancing readers' understanding is important for educators.

A THEORY OF STORIES

In order to understand conflict and rhetorical structure we need to define some terms. We have begun work on a theory which we call the *social interaction model of reading* (Bruce, 1980b). Briefly, a text is viewed as a form of communication, and communication implies social interaction. The social interaction that occurs within a text can be viewed as operating on various levels.

For the purposes of this paper we will gloss over some of the distinctions made in Bruce (1980b), and consider only three of the levels of social interaction (see Figure 1). At level 0 we have the always-present communication that occurs between the real author and the real reader via the written text. Level 1 represents an implied communication between the implied author and the implied reader. Although we speak of the "author of a story", we cannot know whether the author visible to us accurately represents the views of the person by that name. That is, the author we see is really an *implied author* (Booth, 1961). In fact, the implied author is like the real author only to the extent that the real author correctly portrays her or his own beliefs, language, and values. In a similar fashion, the real reader differs from the *implied reader*. At the implied author-implied reader level we will look at rhetorical forms, point of view, and inside view of characters. Finally, level 2 represents character to character social interactions. At the character-to-character level we will examine types of conflict, the response of characters to conflict, and the resolution of conflict.



1. Levels of communication for a single story.

Author-Reader Interactions

Basic rhetorical forms. We have defined six basic rhetorical forms. These forms define the relationship that exists between the implied author and the implied reader and their involvement in the story being told. Full definitions of these forms can be found in Bruce (1980b). For this paper we will define only two of them. A *participant account* is told in the first person. Its point of view is that of the implied author. Unlike the diary form, its implied reader is explicitly not the implied author. In the *observer account* the implied author is not a participant in the actions described. This form is signaled by the use of third person pronouns.

Point of view. For the case of observer accounts, the implied author takes on a perceptual point of view whenever she or he describes an action or a scene. This point of view may be that of a character, or a group of characters, or of no character. A point of view presented in one part of the text may or may not be consistent with that used elsewhere in the text. We have identified four important point of view types: The *consistent single* point of view always, or nearly always, shows events as they might be observed by a character, or a group of characters which stays together throughout the story. The *double* point of view shifts back and forth between two points of view, but otherwise is like "consistent single". The *inconsistent single* point of view generally maintains the perceptual point of view of one character, but occasionally shifts to views unavailable to that character. The *shifting* point of view does not maintain a consistent perceptual point of view.

We have also made a distinction between two types of "main" characters in stories. The *focus character* is the one whom the story is about, that is, the one who is most involved in the principal events of the story. The *point of view character* is the one through whose eyes or over whose shoulder we see the events. Often, of course, a single character fills both of these roles.

Inside view. We define *inside view* to be the portrayal of a character's thoughts and feelings. The implied author may show us anything from the surface of a character to a deep insight into his or her perceptions, emotions, ideas, and so on. The type of inside view presented interacts in intricate ways with the point of view type. For example, a consistent single point of view permits, though it does not require, a deep inside view of one character. A shifting point of view, on the other hand, cannot contain any one deep inside view and usually makes any inside view difficult to maintain.

The scheme presented here calls into question the usual breakdowns of stories into a handful of categories. Consider, for example, the traditional point of view categories (Perrine, 1966): objective, omniscient, and limited omniscient. Each of these captures only a few of the many possibilities outlined above, or else, is ambiguous with respect to which category is intended. The objective structure usually implies limited inside views and no consistent perceptual point of view. However, it is occasionally applied to stories with a consistent point of view but no inside view. Such problems in definition lead to difficulties in analyzing stories. These difficulties in turn were a principal motivation for the development of our taxonomy.

Character to Character Interactions

We saw in the Little Red Riding Hood example that removing conflict from a story also removes its reason for being. Why is conflict so important? There are a number of reasons. First of all, conflicts involve situations or events that are unusual, that are extraordinary, or that in some way alter the status quo. In a sense, they make a story newsworthy. Second, conflicts consist of unknown and uncertain factors which can generate a sense of mystery, curiosity, or suspense and can lead to surprise. We have the feeling that something is going to happen in a conflict, that things are not in a stable state. We wonder, for example, how will the conflict progress? Will it reach a resolution? Third, our interest in the resolution of a conflict relates to our concern for the characters we have come to care about in a story. Their conflicts are important for them, hence for us. Fourth, their conflicts and attempts at resolution can be associated with the conflicts in our experience. People read about conflict partly because conflicts are common in human interactions. Finally, conflicts can be rich, varied, intricate, and complex in the path they take from their initial materialization to their subsequent resolution. Thus, we are intrigued by the possibilities inherent in the path towards resolution. Some of these intricacies are suggested by the definitions to follow.

Conflict Types. Conflict is a situation in which a character or characters are unable to achieve one or more of their desired goals. We define three types of conflict: interpersonal, internal, and environmental. An *interpersonal conflict* exists when two or more characters maintain incompatible goals. An *internal conflict* appears when a single character maintains two or more incompatible goals. An *environmental conflict* exists when a character's goal is hampered by nature, society, or fate. One could study an interpersonal or even an internal conflict from any of the opposing viewpoints.

Responses to Conflicts. In order to study the development of conflict in a story we need to consider how characters respond to conflicts. A *response mode* is a verbalization, a thought, or an action that a character makes subsequent to, and related to the conflict. This implies that the participant, at some level, was aware of his or her involvement in the conflict.

A response mode might or might not be a move predicated to achieve the desired goal. For example, devising a clever solution can be seen as an obvious attempt to attain a goal, whereas engaging in an argument about whether to embark on some course of action toward the goal would be counter-productive. One could also talk about response modes that are believable or reasonable in relationship to the goal, or responses that tend to escalate or de-escalate a conflict. More importantly, one could talk about response modes that are "constructive" in relationship to the goal, i.e., responses that are more likely to bring about the achievement of the goal.

Conflict Resolution. An important element in the structure of story conflict which needs definition is *resolution.* A resolution is a working out of the conflict or an end to the original conflict. From an individual character perspective, it is the relationship between the character and the original goal. That relationship can exist in any of five states. The character could achieve the original goal; partially achieve the goal (essentially a compromise); forsake the goal, willingly and completely giving it up; forsake the original goal but reformulate and adopt a new goal; or fail to achieve the goal and accept the failure, thus abandoning the goal. Finally, in a sixth, non-resolved state, the character could fail to achieve the original goal but not abandon it. He or she would still be embroiled in the original conflict and might well attempt an alternate response mode.

INITIAL STORY SURVEY

The formalizations of rhetorical structure and conflict which we have developed provide a useful framework in which to study children's stories. In order to apply notions made explicit by our analysis we devised a coding form intended for use with primary level children's texts. The form is composed of questions on conflict type, response and resolution modes, rhetorical form, point of view and inside view. In addition, it includes a single metric of conflict complexity we have devised which takes into account such factors as: the number of conflicts per story, the number of different types of conflict, the number of participants involved in story conflicts, the intensity of each conflict, the length of time story conflicts remain in focus, the number of response modes utilized, etc.

We had three main purposes in conducting our initial story survey: (1) determine the prevalence and distribution of the story features illuminated by our analysis in a sample of children's texts, (2) examine the relationship between traditional measures of story complexity, the most well known of these being readability formulas, and our own conflict complexity measure in the sample of texts chosen, and (3) explore relationships that might exist between preference ratings of the stories and the coding of story features such as conflict type, response modes, and inside view. We selected a sample of 32 children's texts composed of 16 upper primary level and 16 lower primary level stories distributed evenly among four groups: popular trade books, random trade books, widely read basal stories, and stories from other educational texts. We then computed the Fog and Spache readability formulas on each of the stories in the sample. Five adult raters were asked to read the 32 stories in the sample and then rank them in order of preference. At a different time, the five raters coded the stories using the form discussed above.

We found there to be 100% agreement among raters that 29 out of the 32 stories exhibited conflict. This finding points to the predominance of conflict in stories, lending support to our claims concerning the importance of story conflict. It is interesting to note that one rater saw conflict in every story. There was 80% agreement among raters that the remaining three stories did not contain conflict. All three of these stories were lower level primary texts. Although these numbers are small, this finding if corroborated in a larger study could raise questions about the traditional emphasis placed on vocabulary and sentence length in beginning readers. Perhaps we are unnecessarily forsaking important text features such as conflict which often lend structure, cohesiveness, excitement and diversity to stories.

Of the three types of conflict discussed above, interpersonal conflict was found to be more frequent and more widespread in our sample. Twenty-eight out of 32 stories exhibited interpersonal conflict, 25 out of 32 stories had environmental conflict and 8 out of the 32 exhibited internal conflict. Thus internal conflict is a relatively infrequent form in the children's stories sampled. When raters were asked to code conflict type for the two most important conflicts in each story, the results were as follows: 59.34% interpersonal conflict, 34.14% environmental conflict, and 6.50% internal conflict. We calculated the distribution of conflict types for the four groups in our sample. Interpersonal conflict was found to be more prevalent in all groups except the popular basal category, where environmental conflicts outnumbered the other two forms. This trend was even more pronounced for the lower level stories. Lower level random trades had the same number of environmental and interpersonal conflicts and no internal conflicts. Lower level popular trades and other educational texts exhibited fewer environmental than interpersonal conflicts and still fewer internal conflicts. However, in the lower level popular basal category environmental conflicts outnumbered interpersonal conflicts by six to one and there were no internal conflicts. For children whose reading exposure is largely limited to school text this somewhat unusual distribution of conflict types and overabundance of environmental conflicts in basal stories may lead to difficulty in understanding conflict forms encountered in reading other texts.

Examination of inside view reported for the 32 stories in our sample reveals an increase in the incidence of high inside view in the upper level stories: 68.7% of the upper level texts in the sample contained high inside views as compared to only 18.7% of the lower primary level texts. This trend was even more pronounced for the basal stories in the sample: 100% of the upper level basal stories displayed high inside views of characters; none of the lower level basals provided similar inside views. This abrupt shift in a key story feature such as inside view is an important finding to investigate further, for it points to a possible explanation for some of the difficulty children encounter in the transition from lower to upper primary level reading.

A second aim of our study was to determine if a relationship existed between the readability scores on our sample which purport to measure story complexity and our own conflict complexity measure. Raters' conflict complexity metrics were transcribed from the coding forms and then averaged across raters for each story. Next, we calculated both the Pearson product-moment correlation coefficient and the Spearman rank correlation coefficient for the Fog readability measures versus the average conflict metrics for the 32 stories. Neither of these coefficients achieved conventional levels of statistical significance (Pearson $r=0.298$; Spearman $r=0.161$). The low correlation suggests that traditional readability measures may be missing important facets of what makes a story complex (see also Bruce & Newman, 1978).

A third aim of our study was to look at the relationship between higher level features of stories and readers' preferences. Using the Spearman rank method we found a statistically significant correlation ($r=.575$, $p<.01$) between the amount of inside view (averaging across five raters' judgments of this feature) and reader preference. The results are still significant when calculated separately for upper level or lower level texts. This result is for adult ratings of inside view and preference. Nevertheless, it suggests the potential importance of inside view as a rhetorical device for creating and maintaining reader interest.

IMPLICATIONS

Our model of author-reader interactions and conflict is one step toward a richer language for discussing stories and for enhancing children's understanding of stories. Its most important contribution may lie in furthering the dialogue between teachers and students regarding the literature which they read. This new language also permits us to examine some other issues more effectively.

One issue is that of defining the readability of texts. The problems children encounter in comprehension may lie not just in the length of sentences or word difficulty, as traditional readability formulas suggest, but in the complexity of the rhetorical structure or the conflicts portrayed in the story. For example, responding to conflict using deception requires a character to view the world from another character's point of view (see the analysis of "Hansel and Gretel" in Bruce & Newman, 1978). Such a shift necessitates inferences that produce a greater overall complexity in the story than there would be if the conflict were confronted directly. Hence, younger children may have difficulty understanding certain stories because they include complex conflict sequences that previously went unnoticed (Bruce, 1980a).

Another issue is reader involvement. If conflict in real life situations has the power to arouse and engage human interest and generate excitement, mystery, curiosity, suspense, and surprise, it is important to study what types and features of conflict in stories could generate the same excitement for a reader. Also, the rhetorical structure/point of view combinations differ in their effect on reader involvement. For instance, an observer account with a shifting point of view and low inside views tends to discourage reader involvement with any one character. We need to know how these effects occur and what the distribution of the rhetorical structure/point of view types is in children's stories.

Another result of these studies could be better criteria for text design and selection. Complaints leveled at some educational texts claim that the conflicts that do exist in stories are monotonous and uninspired (Blom, Waite & Zimet, 1970). On the other hand, fairy tales are said to have survived precisely because they retain familiar conflict patterns in a simple form (Bettelheim, 1976). By applying our model we hope to be able to articulate criteria which will improve the quality of children's stories.

Children's difficulties in understanding texts might be due to mismatched expectations arising from cultural differences. Smitherman (1977) has argued that in black folk tales, to take one example, characters frequently respond to conflict by engaging in clever deception. "The underdog wins by outsmarting his opponent" is a common conflict theme. Also, the black folk tales have a high incidence of commentary by the author, and other distinctive rhetorical structures. Further study of stories from different cultures and subcultures may reveal other distinct patterns. This might indicate the need to diversify the diet of stories given to children.

CONCLUSION

Our studies of children's stories are highlighting features which may account for reader involvement with characters and the implied author, for reader enjoyment, and for difficulty in comprehension. We have been led to consider features which have traditionally been viewed as being in the domain of literary analysis rather than that of reading research. Thus, although these features have direct implications for reading, our examination has taken us far afield from some traditional categories of reading research, such as word and sentence difficulty. We believe it is useful to continue this exploration.

ACKNOWLEDGEMENT

This research was supported by the National Institute of Education under Contact No. US-NIE-C-400-76-0116. We would like to thank Kathy Starr, Andee Rubin, Scott Fertig and Aleida Inglis for help in many segments of the research and Joan Hirschkorn for preparing the manuscript.

REFERENCES

Barber, M. John le Carre: An interrogation. *The New York Times Book Review*, September 25, 1977.

Bettelheim, B. *The uses of enchantment: The meaning and importance of fairy tales.* New York: A. Knopf, Inc., 1976.

Blom, G. E., Waite, R. R., & Zimet, S. G. A motivational content analysis of children's primers. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

Booth, W. C. *The rhetoric of fiction.* Chicago: The University of Chicago Press, 1961.

Bruce, B. Analysis of interacting plans as a guide to the understanding of story structure. *Poetics*, 1980a, *9*, 295-311.

Bruce, B. A social interaction model of reading. *Discourse Processes*, 1980b, in press.

Bruce, B., & Newman, D. Interacting plans. *Cognitive Science*, 1978, *2*, 195-233.

Perrine, L. *Story and structure.* New York: Harcourt, Brace & World, Inc., 1966.

Smitherman, G. *Talkin' and testifyin'.* Boston: Houghton Mifflin, 1977.

ROBERT J. TIERNEY
JAMES MOSENTHAL
University of Illinois

A CRITICAL LOOK AT MACRO-ANALYSES OF TEXT

The purpose of this paper is to consider the use of certain text analysis systems as a means for providing formal representations of text. Specifically, the paper looks at Meyer's system for formally describing expository texts (Meyer, 1975), and at Thornkdyke's grammar for formally describing simple stories (Thorndyke, 1977). As well, Meyer's and Thorndyke's systems are discussed in relation to van Dijk's notions on macro-structures and macro-operators and their incorporation in a theory of text grammar — a theory which means to allow for the formal representation of any text. It is our argument that text analysis is mis-applied when the particular system used claims to formally represent that which is structurally inherent to a text.

The concept of a formal representation of text implies a representation that is rule governed. Such a representation means to formally describe something inherent in the text. This attempt encourages a use and an approach to text structure that treats a text as if it were an absolute. Van Dijk's notion of a text grammar theorizes the legitimacy of formally represented texts while the systems described by Meyer and Thorndyke assume the legitimacy of such representation. Unfortunately, in reading research it is a deceptively satisfying notion to hypothesize a formal description of, or template for, a text which can be used for comparing and evaluating recalls. The fact is, there are problems with the idea of a formal representation of text — at least in so far as macro-analyses are concerned.

Consider a possible formal representation of the plot structure of a simple story proposed by Thorndyke. According to Thorndyke, any story can be generated formally by the categories of the grammar (setting, theme, plot, resolution, etc.) and the rules (never defined explicitly) for generating the plot structure. However, Thorndyke's grammar can be criticized for the arbitrariness with which the grammar's rules are applied to a story. Essentially arbitrary decisions are made concerning the assignment of parts of a story to the categories of the grammar — the decisions are not grammatically determined. In other words, the assignment of parts of a story to a category of the grammar is as much a product of one's world knowledge relative to stories as it is a product of the grammatical description of a story. It could be argued that Thorndyke's grammar is meant only to describe an internalized grammar for stories. However, the idealized description of this grammar, which Thorndyke does mean to describe, still has none of the formal rigor characteristic of the concept of a grammar. In other words, Thorndyke's grammar does not describe any rules or operations, which allow or determine transformations in the grammatical description of a story.

In the case of Meyer, consider the following text and accompanying content structure:

THE GARBAGE COLLECTORS OF THE SEA

The garbage collectors of the sea are the decomposers. Day and night, ocean plants and animals that die, and the body wastes of living animals, slowly drift down to the sea floor. There is a steady rain of such material that builds up on the sea bottom. This is especially true on the continental shelves, where life is rich. It is less true in the desert regions of the deep ocean.

As on the land, different kinds of bacteria also live in the sea. They attack the remains of dead plant and animal tissue and break it down into nutrients. These nutrients are then taken up by plant and animal plankton alike. Among such nutrients are nitrate, phosphate, manganese, silica, and calcium.

As the nutrients are released, they spread around in the water. But they tend to stay near the bottom until some motion of the water stirs them. As you saw earlier, during those seasons when the water is churned up and mixed, the nutrients are brought up to the surface. They may also be brought up by the upwelling action of deep currents. This is especially so along the west coasts of Africa, South America, and North America. Wherever there are regions of upwelling of nutrients, there are rich "fields" of plant plankton, usually during all seasons of the year.

So nutrients are kept circulating endlessly in the oceans, and are used over and over again by plant plankton and whales alike. When a plant or animal breaks down the sugar-fuel it needs for growth, the energy stored by the sugar is used. Some of it goes into building new body parts and some of it is lost as heat. This is not true of nutrients.

Nutrients are not "used up" in that way. For a while, oxygen, carbon, calcium, and other nutrients that a plant or animal takes in become part of the plant or animal. But when the animal or plant dies and when it gives off body wastes, the nutrients are returned to the environment and can be used again and again. (Taken from Ginn Science Program, 1975, Intermediate Level C, p. 395-396.)

Figure 1. The Garbage Collectors of the Sea

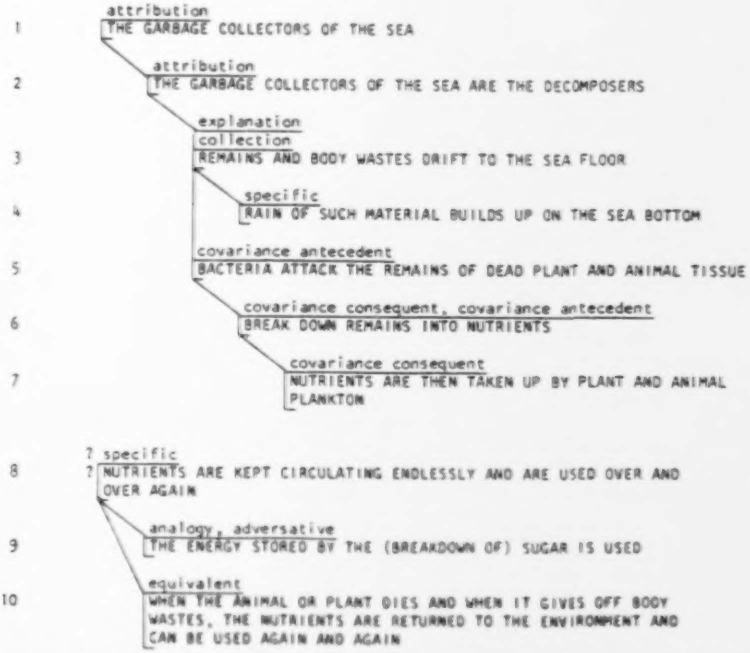

1 attribution
THE GARBAGE COLLECTORS OF THE SEA
2 attribution
THE GARBAGE COLLECTORS OF THE SEA ARE THE DECOMPOSERS
explanation
collection
3 REMAINS AND BODY WASTES DRIFT TO THE SEA FLOOR
specific
4 RAIN OF SUCH MATERIAL BUILDS UP ON THE SEA BOTTOM
covariance antecedent
5 BACTERIA ATTACK THE REMAINS OF DEAD PLANT AND ANIMAL TISSUE
covariance consequent, covariance antecedent
6 BREAK DOWN REMAINS INTO NUTRIENTS
covariance consequent
7 NUTRIENTS ARE THEN TAKEN UP BY PLANT AND ANIMAL PLANKTON
? specific
8 ? NUTRIENTS ARE KEPT CIRCULATING ENDLESSLY AND ARE USED OVER AND OVER AGAIN
analogy, adversative
9 THE ENERGY STORED BY THE (BREAKDOWN OF) SUGAR IS USED
equivalent
10 WHEN THE ANIMAL OR PLANT DIES AND WHEN IT GIVES OFF BODY WASTES, THE NUTRIENTS ARE RETURNED TO THE ENVIRONMENT AND CAN BE USED AGAIN AND AGAIN

Figure 2. A Meyer Analysis of the Structural Properties of 'Garbage Collectors of the Sea'

In Meyer's content structure the underlined terms written in lower case letters represent rhetorical predicates describing the relationship between parts of a text. Such a rhetorical predicate as covariance antecedent/consequent describes a relationship of causality or allowability between parts of the text. Left most entries describe superordinate propositions of the text (macro-propositions). Overall, the content structure means to present a formal representation of that structuring of ideas which is inherent to the text and which determines comprehension of the text.

However, the arbitrariness of the determination of rhetorical predicates and the actual propositions is evident when discussing entries eight through ten. These entries, dealing with the topic of nutrients, should appear between entries six and seven — subordinate to those entries. But, it is apparent from reading the text that entry eight should appear as a left most entry since the topic of nutrients dominates most of the passage. This is the case although the text itself does not initially prepare for this dominance and does not allow for any formal 'vertical' tie between the left most entries one and eight.

This problem of arbitrariness is not just a characteristic of Meyer's and Thorndyke's text analysis systems which mean to formally represent text; but it is also characteristic of van Dijk's more formal approach. In his article, "Text grammar and text logic" (van Dijk, 1973), van Dijk presents a text logic describing the micro-structure of a text. More recently (van Dijk, 1977a, 1977b), he implies that with macro-operators one can provide a formal proof for a macro-structure of a text. He describes four macro-operations: construction, generalization, integration, and deletion. The extent to which these macro-operations are formal operations is evident from van Dijk's definition of the four operations. Consider the definition of the integration operation:

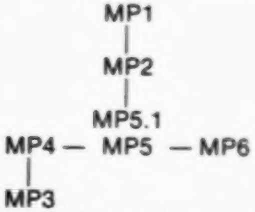
> Given a sequence of propositions $\Sigma = (p_1, p_2, \ldots, p_n)$, $n \geqslant 2$, if there is a proposition p_i such that for each proposition p_j, $V(p_j)$ is either a normal condition, a normal component, or a normal consequence of $V(p_i)$, then we may substitute p_i for Σ. (van Dijk, 1977a, p. 13)

In this definition $V(p_n)$ denotes the fact or truth of p_n relative to the 'world' described by the text. In practice, elements of the formal description of the micro-structure of a text are operated on by the integration operator (or one of the other operators). Through progressive applications of the operators progressively more general levels of a macro-structure are developed.

Consider the informal application of van Dijk's text logic to the first paragraph of 'Garbage Collectors of the Sea.' The micro-structure of this text might consist of the following list of micro-propositions:

1. (there are) garbage collectors of the sea (from the title)
2. the garbage collectors of the sea are the decomposers
3. day and night, ocean plants and animals die
4. and slowly drift down to the sea floor
5. day and night the body wastes of living animals slowly drift to the sea floor
6. there is a steady rain of such material
7. that builds up on the sea bottom
8. this is especially true on the continental shelves
9. where life is rich
10. it is less true in the desert regions of the deep ocean

Such a list could be given for the entire text. If the macro-operators construction, generalization, integration, and deletion are informally applied to the list of micro-propositions the following diagram of a macro-structure for the text might develop:



where the macro-propositions are:

MP1 there are garbage collectors of the sea
MP2 the garbage collectors of the sea are the decomposers
MP3 the remains of ocean plants and animals, and the body wastes of living animals build up on the sea floor
MP4 organic material builds up on the sea floor
MP5.1 the decomposers are the bacteria
MP6 the nutrients are used over and over again

It is apparent that this macro-structure represents a summary of the text. The point of setting up this macro-structure is to indicate that, even with a formally determined micro-structure and the formal application of macro-operators, the arbitrariness of the decisions necessary for the generation of certain macro-operators such as MP5.0 remains evident. For example, no informal use of macro-operators (assumed or stated) in a macro-analysis of a text, can absolutely determine such a macro-proposition as MP5.1. Yet, the importance of MP5.1 is apparent in the macro-structure diagram. Without it no tie is made between macro-propositions one and two, and the macro-propositions describing the cycling of nutrients. It is a problem in the macro-analysis of the text that such a strong inference must be made arbitrarily from world knowledge in order to relate explicitly all aspects of the text in the macro-analysis.

Despite the shortcomings of van Dijk's macro-analysis the general notion of a text logic as he proposes would seem legitimate, for the attempt to set up such a logic is a necessary pre-requisite to deriving a resultant macro-structure depicting a formal representation of text. Without macro-operators and a formal micro-structure the determination of a macro-structure (that which is structurally inherent to a text) would seem an impossibility. In other words without both macro-operators and a micro-structure the claim that the macro-structure of a text is formal and rule governed seems theoretically misconceived. Unfortunately, neither Meyer nor Thorndyke describe formally determined micro-structures or a formal procedure for the macro-analysis of a text. And, in its present form, van Dijk's text logic and macro-operators seem inadequate (see Kiefer, 1977; Morgan & Sellner, in press, for additional criticism).

With these limitations of macro-analyses in mind, it is appropriate, in the field of reading research, that the goals of text-analysis be re-evaluated—especially such goals as formulated at the National Institute of Education's Conference on Studies in Reading, 1975. In the report of that conference, it was stated that

> a necessary pre-requisite to research on the structural properties of text and their interaction with the processing characteristics of the reader is the development of a linguistic model of discourse structure. (*NIE Conference on studies in reading*, 1975)

If it is not an impossible attempt, it is at least, at this time, an inadequate analogy to suppose that the structural properties of texts can be parsed as if they have the structural integrity of a sentence. From a more practical perspective, it is not that macro-analyses of texts are totally inappropriate in experimental research — it is that they are appropriate only as one possible tool reflecting, not defining, processes of text comprehension. In the case of macro-operations, it is best, when dealing with text, to use the concept of macro-operators as descriptors for cognitive processes representing general principles of perception, reasoning, and world knowledge. Researchers intent on using macro-analyses of text should be aware of the theoretical fallibility of their system for analyzing texts. Macro-analyses of text do not, as the systems of Meyer and Thorndyke suggest, describe inherent, absolute structural properties of text. At the very least, researchers should provide their audiences with the actual macro-analyses used and should comment on the generalizability of their macro-models as well as comment on the extent to which they are arbitrarily derived.

REFERENCES

Gallant, R. A. The garbage collectors of the sea. In I. Asimov & R. A. Gallant, *Ginn Science Program* (Intermediate Level C). Lexington, Mass: Ginn & Co., 1975.

Kiefer, F. Review of S. J. Petofi & H. Rieser (Eds.), Studies in text grammar. *Journal of Pragmatics*, 1977, *1*, 177-192.

Meyer, B. *The organization of prose and its effect on memory*. Amsterdam: North-Holland, 1975.

Morgan, J. L., & Sellner, M. B. Discourse and linguistic theory. To appear in R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Erlbaum, in press.

NIE Conference on studies in reading (Vol 2): *The structure of language. Washington, D.C.: National Institute of Education, 1975.*

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

van Dijk, T. A. Text grammar and text logic. In J. S. Petofi & H. Rieser (Eds.), *Studies in text grammar*. Dordrecht-Holland: D. Reidel, 1973.

van Dijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Erlbaum, 1977. (a)

van Dijk, T. A. *Text and context: Explorations in the semantics and pragmatics of discourse*. New York: Longman, 1977 (b).

LYNNE REDER
Carnegie-Mellon University

A SHORT SURVEY OF THE RESEARCH ON TEXT AND STORY COMPREHENSION AND RETENTION

INTRODUCTION

In the last half decade, an interest in studying prose comprehension and retention has been rekindled among cognitive psychologists. The psychologist generally credited with first studying memory for stories was Bartlett (1932). His early research is still frequently cited and still influences the field. In his experiments, subjects read stories from different cultures which they later recalled. The main finding was that recall was inaccurate; elements of the story which were poorly understood for cultural reasons were distorted in recall to conform better with his subjects' knowledge. He considered his results incompatible with the notion of "fixed lifeless traces" and instead posited that remembering is a *constructive* process based on general impressions and attitudes.

For approximately 30 years, since Bartlett's seminal work, psychologists left the investigation of reading to researchers in education. Instead, psychologists looked at memory for lists of nonsense syllables, words and unrelated sentences. One reason psychologists may have shied away from studying prose is the problems inherent in using large units of analysis. For example, it is difficult to determine the similarity and differences among passages, which in turn makes comparisons across experiments that use prose difficult. There have been two main dependent measures employed in the study of prose comprehension and retention: question answering and free recall of the passage. Anderson (1972) has argued that proper selection of questions intended to measure comprehension is not a trivial task. And, just as it is difficult to evaluate the similarity of prose passages, it is difficult to score recall of prose in a meaningful way to estimate how much has been retained.

The new willingness to study prose comprehension, despite its problems of imprecision is due to several factors. First of all, some investigators have lost faith in the idea that results from simple laboratory experiments will generalize to natural materials. Second, it is now easier to conduct research and develop theories related to reading than it was a few years ago, due to the developments in computer science and linguistics.

Computers are an aid to the study of prose in two ways. Paragraphs can now be presented to the subject "on-line", each sentence displayed on a computer terminal for only the time a specific subject needs to read it. Reading times can be recorded, questions can be asked at any desired point, speed/accuracy in answering the questions can be manipulated and so on. Some of the most promising work uses computers to also monitor eye fixations (Carpenter & Just, 1977; Just & Carpenter, 1978). Computers are also an aid to theoretical development as they allow simulation of prose processing. It would be very difficult to develop a model of prose processing in much detail without implementing it on a computer as the complexity is not conceptually tractable. Another reason theorists have returned to prose processing is the concomitant development of text grammars in linguistics, (e.g., Lakoff, 1972). Experimenters and simulators use theoretical notions from linguistics to suggest experiments or ways to simulate.

There are other groups of investigators apart from cognitive psychologists who have been studying prose. Researchers in the field of education have been exploring this issue for many years (e.g., Ausubel, Frase, Gagne, Rothkopf). Memory theorists who have uncovered a number of principles using the "verbal learning materials" (e.g., lists of nonsense syllables or words) have investigated to what extent the principles generalize to connected discourse.

One of the problems, of course, in studying prose comprehension is the need to pare down the morass of complexities associated with prose. Thus far, the strategy to deal with the enormous complexity seems to be to partition the problem of prose understanding into subproblems and only focus on one or two aspects, e.g., the structure and grammars of stories, how the "content" (words and topics) affects recall, how orienting tasks affect comprehension, or the role of knowledge bundles—scripts or frames—in understanding. Hopefully, this problem can be partitioned into different aspects, but unless investigators are aware of all aspects of the problem, little success is likely.

The first section of this paper will review some of the research that has focused on applied questions of interest to education. The results from some of this research indicate the need for a more sophisticated understanding of text structures and internal representations of text. A review of current work on text grammars is given, followed by a discussion of work on processing models.

INVESTIGATIONS OF FACTORS THAT AFFECT AMOUNT OF RECALL

Focus on Instructionally Relevant Experiments

An important application of research dealing with prose processing is how to improve people's ability to read. Various techniques have been tried to improve the student's comprehension and retention of reading material. Ausubel (1960) introduced the notion of "advance organizers". He thought that adjunct aids which give the reader a preview of the content of the passage would improve the reader's organization of the material, thus leading to better comprehension and retention. The advance organizers should activate or develop appropriate cognitive structures to which subjects could anchor the incoming ideas. He predicted that prior exposure to subsuming concepts or "anchors" would facilitate retention. Ausubel and his associates (e.g., Ausubel & Fitzgerald, 1962) have found data to be consistent with these notions to some extent; they admit, however, that an advance organizer's usefulness interacts with many other variables such as previous knowledge and verbal ability.

The investigation of the use of adjunct aids in the comprehension of prose has been continued by a number of investigators. One such aid that has been explored thoroughly is the effect of questions on *subsequent* test performance (e.g., Anderson & Biddle, 1975; Frase, 1967, 1968, 1971, 1972, 1975; McGaw & Grotelueschen, 1972; Rothkopf, 1966, 1972; Rothkopf & Bisbicos, 1967; Watts & Anderson, 1971; see Anderson & Biddle for an extensive review). The typical dependent measure is how well a subject can answer a set of questions about a passage (call these *test questions*) as a function of having been asked questions about the passage earlier, (call these *priming questions*). Examples of the independent variables are whether the priming questions are asked before the subject reads the passage, during the passage or after the passage; whether the priming questions are general or specific; whether the subject is given feedback on the priming questions.

The majority of the studies in the field have found that subjects do better at answering a test question if they have been asked that *same* question earlier. The improvement obtains regardless of whether the priming question was given before the passage was read or afterwards and whether or not the subject was given feedback about his answer. Performance is best if the question was asked before the passage was read. In the case where priming questions ask about information different from that to be tested, the findings are mixed: many studies have found that subjects do better than control subjects (asked no questions) on the non-primed test questions but only when the priming questions were asked *after* the passage was read. If the priming questions were asked prior to reading the text, performance is occasionally worse for experimental subjects than it is for controls on non-primed material.

Priming questions provide a focus that tells the subject what aspects of the text are important. However, providing a focus can not be the entire explanation of the improvement due to priming questions. The location of priming questions is also important. Frase (1967) found that asking a question after each paragraph is superior to asking all the priming questions at the end of the passage. Frase (1967) also found that asking a priming question after two paragraphs was better than after one or four. If the effect of priming questions were merely to provide a focus then the position of questions should not matter.

It seems that priming questions do more than provide focus; they force the subject to process the text in a certain way. This is clear from results of experiments where critical aspects of the text were directly highlighted (in a box) at study. Bruning's (1968) study found that this method of highlighting was not nearly as effective as forcing the subject to answer a question that caused that material to be reviewed in order to answer it. The conclusion that providing information may not be as effective as forcing the subject to retrieve it himself has been encountered before. Bobrow and Bower (1969) found that providing the mnemonic to relate a pair of words was not nearly as effective as asking subjects to provide one themselves.

Frase (1967) suggested that priming questions cause the subject to review the relevant aspects of the passage. The review process probably involves more than merely stating the critical information in the question or answer since highlighting the critical information was so much less effective. Indeed, McGaw and Grotelueschen (1972) found that the information does not need to be directly tested. In their study, when a question reminded a subject of information present in the text, without either stating the information in the question or demanding it as an answer, subjects later recalled that fact better than a comparable fact unrelated to the priming question. The reason that priming questions in studies mentioned earlier for the most part only helped some questions and not new test questions is probably due to the nature of the overlap between questions. In fact, Reder (1976, 1979) and Watts and Anderson (1971) both conducted experiments in which the priming questions both helped some new test questions and not others, depending on the type of questions or the type of relation between priming and test questions.

It seems clear that merely asking a question will not produce improved performance. The question must force the subject to process relevant aspects of the text in "useful" ways. Watts and Anderson (1971), for example, found that subjects do better on a post-test about passages they read if they had been asked immediately after reading each passage a question that forced them to "integrate" the material (i.e., apply a principle in the passage to a new example). On the other hand, subjects do even worse on the post-test if asked a low-level question (name the scientist associated with the principle) than asked no question at all.

One explanation for this result involves the notion of generating useful elaborations. This notion is that the more extra processing one does that results in additional, related or redundant propositions, the better will be memory for the material processed. The importance of relevant elaborations or embellishments to memory has been argued elsewhere (Reder, 1976, 1979) and will be discussed in more depth later. Elaborations provide redundancy in the memory structure. Redundancy can be viewed as a safe-guard against forgetting and an aid to fast retrieval. In the Watts and Anderson study, when subjects were asked no questions they probably had more time to generate useful elaborations than when they were forced to generate the name of the scientist which has nothing to do with the point of the principle. In other words, it is not enough to force people to think about the relevant information, they must generate semantically useful elaborations; if they think about aspects of the text that seem trivial and unimportant (e.g., when reading about how electricity works, thinking about the cost of electricity today), then they will not be able to answer questions about aspects of the passage that are deemed important.

Rothkopf (1972) conducted a study whose results are consistent with this view. In this study subjects read material at their own rate and they would read slower those passages for which they were given priming questions. They probably were attending (thinking, elaborating) more to the relevant information in the primed passages. Anderson and Biddle (1975) also produced results consistent with the notion of greater amount of attention producing better performance. In their experiments subjects were asked a priming question after reading the passage that could either test verbatim memory or gist memory. They found that a priming test of verbatim memory produced better subsequent performance on either a verbatim or paraphrase test than did a previous paraphrase memory test. If one assumes that one makes multiple redundant elaborations related to a statement during study, finding anyone will suffice for a gist recall test. However, to decide if something was said in that exact form requires a much more thorough search of one's memory representation. Perhaps verbatim search, being more time consuming, allows more elaborations to be rehearsed, (strengthened) and more new ones to be generated.

Summary. The research discussed thus far has dealt principally with how to improve retention of prose material. The major thrust of this research has been on the effects of questions on text-comprehension. The independent measures are usually whether or not priming questions or "orienting directions" are given, the location of these questions—either before, during or after reading the text and the nature of the overlap of the priming questions or directions with the criterion questions. Several general conclusions can be summarized from this research:

(1) People perform better on criterion tests if previously given orienting instructions or priming questions regarding the text.

(2) The extent of benefit depends in part on the overlap between priming questions and the criterion questions. For example, Frase (1975) gave subjects learning directions such as "pay attention to the number of rotations of Mars around the sun that occur in one year." Control subjects were given no directions: Experimental subjects did slightly worse than control subjects on questions irrelevant to the directions (52% correct vs. 57% correct). However, the experimental subjects did much better than controls on questions relevant to the orienting instructions (76% vs. 57%). So, *overall*, there is a clear advantage for asking subjects priming questions.

(3) There are three positions that are contrasted for giving orienting directions or asking questions: (a) before reading the passage; (b) during the passage, e.g., after each paragraph; and (c) immediately after reading the entire passage. Performance tends to be slightly worse

on *relevant* questions when priming questions come after reading the passage. However, performance on incidental questions is greatly improved if these orienting directions are not given until after the passage has been read. Presumably, subjects will search out the answers to the primed questions in the text and totally ignore all other information when given these directions in advance; if given afterwards, only the review and rehearsal processes are affected. A qualification should be made, however, such that the orienting directions must force the reader to process relevant aspects of the text in "*useful* ways" not just attend to the material.

In sum, forcing readers to attend to those aspects of the passage that are deemed important results in better retention of that information. Indeed, Reder and Anderson (1980) found that subjects can better answer questions about a chapter if they read a summary rather than the original.

Research Using Additional Context as External Aids

The research of Bransford and Dooling illustrate the importance of a relevant conceptual structure within which to embed new information for long-term retention of the input. Like the research discussed in the last section, this work explores the usefulness of adjunct aids. In this work, the adjunct aids consist of pictures, titles, topic phrases, etc. rather than questions. The goal of the research has been less concerned with techniques to improve recall (although recall has certainly been affected by the use of the aids), and has been more concerned with what affects comprehension broadly speaking. In many experiments the passages used were basically incomprehensible without such aids, while much clearer when aids were given in advance. The effectiveness of the aids has been explained in terms of Bartlett's notions of schemata. Bransford and Johnson (1973) and Bransford and McCarrell (1975) found that comprehension and ability to remember a passage were greatly improved when a descriptive title or picture was given prior to the passage. Without the title or picture, the passages were too abstract; the referents needed to understand the passage could only be inferred when the topic was given (in one case) and from the situation depicted by the picture (in the other case). In other words, comprehension depended on a schema or conceptual framework supplied by the experimenter to make the referents of the passage clear. Dooling and Lachman (1971) also presented passages that were much easier to recall if a topic (e.g., "The first space trip to the moon") was given before reading the passage. Giving the topic after having read the passage produced no better recall than not giving it at all (Bransford & Johnson, 1973; Dooling & Mullet, 1973). This indicates, I think, that if one cannot fit the input into a conceptual structure *immediately*, the phrases will not be transferred into long-term memory. Thus, the topic is of no use in the after condition as the story is already forgotten. It is also the case that providing the subject with the wrong topic before reading an abstract passage makes recall worse than having no topic (Bransford & Johnson, 1973). Perhaps this is because the reader wastes processing time trying to integrate the input into the wrong structure or else creates inappropriate elaborations. Had he been given no structure, the reader stood a better chance of deriving an appropriate one. These findings are consistent with Ausubel's suggestion that the reader profits from being given prior structure on which to map the incoming material.

Bransford and Franks (1971) argue that in "normal passages", when the referents are clear, a schematic representation of the input is constructed and only that schema is stored in memory. They showed that subjects were poor at discriminating exact sentences they heard from other sentences that were also derived from the "schema". Bransford and Johnson also showed that if the passages are concrete, (i.e., the referents are clear), prior knowledge of the topic does not aid comprehension or recall. This finding is not inconsistent with Ausubel's assertion that prior structure helps comprehension. He had found that an advance organizer's usefulness interacts with the reader's knowledge and the subject matter. In some "concrete" passages for which the reader knows little about the topic, prior exposure to the structure would undoubtedly improve comprehension and retention. Perhaps the term "concrete" should be defined with respect to each individual's semantic memory since various referents may not seem clear if the reader is unfamiliar with the topic.

Sulin and Dooling (1974) demonstrated the influence of prior knowledge in a different manner. They argued for the importance of *thematic* effects in retention of prose. When a subject read a passage that was said to be about a famous person, e.g., Hitler, they were less confident about rejecting recognition foils that were dominant traits of the famous person than less thematically related foils. When subjects were told that the passage read was about a fictitious character, they rejected all foils with about the same confidence. Sulin and Dooling were able to find in one of their two experiments that the "thematicity" effect increased with delay, as predicted by Bartlett. However, most of these results are open to potential artifacts: The effect could be merely a response bias rather than a difference in encoding. To demonstrate that it is not a response bias, they should have contrasted their results with those from subjects who were not told the identity of the character (Hitler vs. Gerald Martin) until the time of test. Unlike a response bias model, a "schema" model would not predict a replication of their results in the control condition. This is because according to the schema theory, it is during input that one reorganizes, interprets, etc., not at test. It should be pointed out that subjects rarely false-alarmed (i.e., said the foils had been presented); they just gave lower confidence ratings about related foils being new than about less related foils.

More recently, Dooling and Christiaansen (1977) performed the needed control experiment. The subject could be told the identity either before reading the passage, immediately after reading the passage or immediately before test one week later. They found that subjects made many false alarms at a week delay regardless of when they learned the "identity" of the famous character. The percentage of thematic false alarms was greatest when subjects were told immediately after reading the story. Neither the simple response-bias model nor the simple schema model predict the obtained pattern of data, that most thematic false alarms occur when subjects are informed right after reading the story. However, a reconstructive model plus a type of response bias model together could handle the data. The model would posit that if the subject learns the identity of the fictitious character immediately after reading, then the subject will go back over the story, stored in memory, reworking it to make it consistent with the "true" identity. This causes more thematically related inferences to be stored erroneously than when the subject knows while reading the story that the passage is about a famous character. A response bias must also be operating (a pre-disposition to respond in a manner consistent with knowledge about the famous character), to account for the greater number of false alarms to thematic foils when told the character's identity immediately before test.

One interpretation of Sulin and Dooling's result is that subjects need not store the related facts in long-term memory since the result is partially a response bias. This view is supported by data of Frase (1970, 1971). When subjects were asked to rate the validity of statements with respect to a passage, no difference in ratings of inferences and assertions were found. Frase did find differences in recall, however. Very few inferences tended to be produced. This might be because the inferences were not made until subjects were asked to judge their validity; or they might have been made, but tagged as inferences and not considered appropriate for recall. The former seems more likely because when subjects were forced to draw the inference to answer another question, the incidence of the inference's recall increased. Of course, some of both may be going on; the more stored inferences there are, the greater the probability that some tags will be lost.

Conclusion. In conclusion, the findings of Bransford and Dooling discussed earlier, like Bartlett's, provide evidence that context and general knowledge affect memory for prose. They go beyond Bartlett's work by showing when "distortion" of memory can occur (i.e., immediately after reading as well as during reading) and when the requisite knowledge is needed for comprehension and retention (i.e., it must *precede* reading). Like Bartlett, however, the investigators offer few insights as to what the structure of schemas might be like or the mechanisms that allow schemas to affect comprehension. Anderson and Biddle (1975) remark with respect to the questioning literature reviewed in the last section that "We do not need another demonstration that adjunct questions 'work' . . . (we need to know) . . . why they work . . . (p.108)". The same basic comment could be made with respect to these demonstrations.

REPRESENTATIONS OF TEXT STRUCTURE

Many theorists concerned with finding an adequate representation of a prose passage maintain that doing so is a necessary prerequisite to other research on prose comprehension. Meyer (1975) has argued that structural variables can influence the learning and retention of items in a word list and that structure probably also affects retention of prose passages. She points out that one can ameliorate problems such as measuring the similarity and differences among passages and scoring the recall protocol from a passage by the use of a structural analysis.

There have been a number of approaches to the problem of specifying the representation of a text or a story. The first sub-section to follow concerns representations of the relations found in a text or a story. The second part describes theories that have rules for generating a story. Each representational theory is concerned with either expository text or stories, rarely both.

Representations of Relations in a Passage

Frederiksen. Frederiksen's (1972) model bears resemblance to that of Dawes (1964, 1966) and Crothers (1972) in his concern with logical relations. His semantic structure graph uses both set relations and subject transformations (i.e., implication relations). Frederiksen performed an experiment using Dawes' Circle Island story and he has analyzed it in a number of ways (Frederiksen, 1972, 1975a, 1975b). Subjects read the story four times: two of the three groups recalled after each reading and the other group was given a surprise recall after the last reading of the story. They all recalled once more a week later. The uninformed group and one of the two informed groups also had an orienting task of a problem solving nature, related to the passage, to work on after each reading. The dependent measure was recall and Frederiksen developed a scoring system that noted whether concepts were overspecified (pseudodiscriminated), e.g., colonel for officer, or underspecified (overgeneralized) e.g., military man for officer. He looked at how orienting tasks affected the specifications of both concepts and set relations and how this recall changed as a function of successive recall trials.

The group given a surprise test recalled less than the informed groups. This group and the other problem solving group produced more elaborations (statements not derivable "inferentially" from the passage), overgeneralizations and inferred relations than the group just asked to recall. Both informed recall groups gave more "veridical" output than the uninformed group, and the two informed groups did not differ in the amount of veridical recall. The non-problem solving group gave more pseudo-discriminations. Frederiksen believed that the problem solvers had a greater processing load and would therefore have to encode the input more simply. He argued that the inferences produced in recall are products of simpler coding of the input. With these two assumptions he explained the greater amount of inferential recall of the problem solvers. However, this argument seems wrong for several reasons: First, drawing inferences from the input should require more processing, not less. Second, the amount of veridical recall did not negatively correlate with inferential or derived recall. Therefore inferences could not result from a "greater processing load" in the problem solving condition. Perhaps instead, "semantically relevant" memory errors occur more frequently for subjects given a meaningful orienting task to direct their prose processing because meaningful processing involves building a representation of the gist and ignoring the exact input phrases. A greater processing load theory would imply, on the other hand, that these subjects should lose more of the information irrelevant to the task. There was no evidence of this. One explanation of why problem solvers produced more elaborations not derivable from the story is that they import more world knowledge in order to solve the problems. These imported facts needed for solving problems may become confused with the input.

In analyzing the errors over successive recalls, Frederiksen (1975a, 1975b) found more inferred concepts and more overgeneralizations on each output trial, but successive recalls lead to fewer elaborations. He considers this evidence that generalizations occur during input (reading the test) and elaborations are generated during recall. Presumably elaborations are edited during the input mechanism and generalizations are not, because generalizations are not corrected on successive input trials. There is another interpretation for this result, however. The more often one recalls, the more boring the procedure becomes and hence, the more shorthand expressions one looks for. The group that only recalled once may have put down more elaborations because they had little veridical recall and wanted to say as much as possible. The groups who were then recalling for the fourth time may have omitted the elaborations to get through the task as quickly as possible. The inferences and overgeneralizations may have the property of making the output shorter; any shortcut that expressed the same idea would be desirable.

To test this alternative explanation, however, a control group would be needed that Frederiksen did not provide: a group that read the story once, but recalled four times. His theory would predict no additional inferences after the first recall, but more elaborations on each successive output. The other view predicts the opposite outcome.

One problem with Frederiksen's theoretical claim is that they are all based on one study using one story, Circle Island. Dawes wrote Circle Island to investigate errors of logical relations. Frederiksen's representation for stories, like Crothers' will probably not fit most stories. To my knowledge, this representation has not been tried out on a variety of passages.

Frederiksen's (1975c) more recent model has become more detailed and takes into account far more aspects of language than his earlier representations. His representation involves two graphs, the semantic structure graph and the logical structure graph. The first is reminiscent of Fillmore's (1968) case grammar for sentences. He diagrams stative relations, case relations, attributive relations, etc. into semantic hierarchies or networks, making fine, semantic feature distinctions. The logical structure graph is not hierarchical; propositions from the semantic structure graph are nodes connected together by the logical, causal and algebraic relations. The logical relations and causal relations include conjunction, disjunction, logical implication, and material biconditional. The algebraic relations are order (transitive) relations and proximity relations. He can also represent relative time and location.

Frederiksen's representation is one of the few to include inferences. That makes it a more realistic representation of what the long-term memory structure might be like as a result of comprehending a passage. Unfortunately, his representations seem complex at the wrong "level" of understanding of a passage; his system seems devised for representing the relations within a sentence, not among sentences. The network has capacities for connecting propositions referentially and temporally, but apart from these, there is little in his system that seems concerned with the text level as opposed to the sentence level. Textual aspects which he does not represent include the overall organization of the text of the plot and inferences and elaborations that connect lines of input.

Kintsch. Kintsch (1974) has also developed a formal representation for texts and has performed numerous experiments to test the validity of this theory. The fundamental unit in his analysis is the proposition. A text is represented as an ordered list of propositions. He avoids network representations due to their impracticality for passages of any length or complexity. Networks are messy for single propositions, and they rapidly become unwieldy when expressing multiple propositions. A sentence may consist of one or more simple propositions, each proposition consisting in turn of a relation and a set of arguments. The representation has the flavor of predicate calculus. Verbs and adjectives are relations (see attributes), and nouns and other propositions are its arguments. Propositions are connected together by means of a repetition rule, in which one proposition is referred to by a "superordinate" proposition that contains this subordinate as an argument. This forms a hierarchical structure. Propositions are also connected by sharing the same arguments or concepts, e.g., an individual of a story can be in multiple propositons. The first proposition that uses the shared argument is considered the superordinate of the proposition that contains the repeated argument. Propositions can also be connected by an ampersand (&) when they share a common argument. For example, Kintsch (1974, p. 63) represents "John broke his leg yesterday" as: (BREAK, JOHN, LEG) = α & (TIME, α , YESTERDAY). Kintsch claims his hierarchical organization of text has predictive value for recall.

Kintsch reports numerous experiments in support of his theory (e.g., Kintsch & Keenan, 1973; Kintsch, McKoon & Keenan, 1974; Kintsch & Monk, 1972). He admits that these experiments support a class of representations, of which his is one; some studies though (e.g., Kintsch, 1974, ch. 7) were designed to distinguish between his propositional representation and other extant proposition representations. Below a few of his studies will be described in varying detail.

Kintsch and Monk (1972; Kintsch, 1974, ch. 5) argued that ideas in a text are represented in propositional format which does not depend on the surface structure complexity of the input. Two versions of a passage were constructed, one simple in structure and one complex although both were presumed to convey the same ideas. Subjects read one of the two versions. Reading time was of interest as was time to judge that an inference from the story was valid. They found that reading times were longer for complex passages, but that latencies to judge inferences did not vary as a function of complexity. These findings are consistent with the notion that it takes longer to comprehend complex statements and encode them into a propositional format; once stored in their canonical representation, however, the time to make the inference should not reflect surface structure. Kintsch also suggests that recall can be predicted by level in the hierarchy, superordinate propositions being recalled more frequently. Kintsch and Keenan (1973) found this trend. McKoon (1977) found subjects are faster and more accurate to verify topic information than details and that the discrepancy between topic and details grows with delay.

In some respects, Kintsch's (1974) theory is rather similar to Frederiksen's (1975c) theory. Frederiksen gives more emphasis to a low-level (sentence) representation while Kintsch's hierarchical structure seems somewhat better suited to text organization. Certainly a linear notation is more tractable for text representations than is a network structure, although formally, there is no difference. Both theorists include mechanisms for cross-proposition inferences. However, they both seem more concerned with representing sentences within a passage than the ideas of the passage itself, i.e., the inferences in both systems seem to be narrow in scope, concerned primarily with such problems as causation and reference.

The inferences one would like to see in their representations or models for comprehension are those derived from knowledge about the passage as a whole, e.g., inferences about characters' motivation, goals, and responses to events. Elaborations about spatial locations, instantiation of physical details of people and objects are probably made during comprehension, too. Perhaps Kintsch and Frederiksen consider the elaborations suggested above to be idiosyncratic, and therefore, impossible to represent.

Meyer. Meyer (1975), like Kintsch, has a hierarchical rather than a network representation.

Her theory adapted from Grimes (1975) has a greater resemblance to Fillmore's (1968) grammar than do other stories. The similarity between her model and Fillmore's is in the emphasis of the role or case terms to represent the relation between the verb (predicate) and the other words (arguments) in the sentence. (Use of Fillmore's case grammar has also been adopted to a lesser extent by Rumelhart, Lindsay and Norman, 1972; Kintsch, 1972, 1974; and Frederiksen, 1972, 1975c). Meyer lists nine cases or roles quite similar to Fillmore's, which she calls "lexical predicates". Unlike Fillmore, (but very much like Grimes), she has "rhetorical predicates" at a level above the lexical predicates. These higher level predicates give prose its organization. Neither Frederiksen's nor Kintsch's representations (discussed above) have such a well-developed, higher-level organization.

The representation Meyer uses resembles an outline of a passage. Propositions are indented according to their importance. Unlike an outline, however, every idea is included and relationships are indicated between branches at the same level and at differing levels. These relations are the rhetorical predicates and they fall into three categories: (a) paratactic relations, which have at least two arguments of equal weight (e.g., RESPONSE: problems and solution); (b) hypotactic relations in which the arguments are unequal (e.g., SPECIFIC: problem and the details of the problem; and (c) neutral relations in which a collection of arguments can have equal weight or not, (e.g., ATTRIBUTION).

Meyer has attempted to validate her representation empirically with a recall experiment. The experiment was designed to show that the probability of recall of a sentence was a function of its "height" in the content structure (i.e., its importance indicated by degree of indentation). Meyer constructed two passages with the identical target paragraph embedded in each. The paragraph was placed high in content structure in one passage and low in the other, but its serial position was the same. She found that the two passages as a whole were recalled equally well but that the recall of the target paragraph differed as a function of "height". The difference in recall between the two paragraphs increased with a week's delay.

After the free recall task at a week's delay, subjects were given a cued recall test. Cued recall in her experiment meant giving subjects all the content words from the passage in any unsystematic (scrambled) order and asking them to arrange the words in the correct order. Even in this task, the advantage of the paragraph high in the content structure was maintained over the paragraph low in the content structure. From these results, Meyer concluded that the difference in memory representation for "low" and "high" paragraphs can not just be a function of differential accessibility. She argues that if the paragraphs were stored equally well, any differences due to accessibility would be avoided by the cued recall methodology. Meyer's task, however, seems sufficiently different from the standard cued recall task that differential accessibility still seems a possible explanation. People have great difficulty solving anagrams, yet few would doubt that these puzzle solvers have the word (whose letters are presented in scrambled order) stored in memory. It may well be in her task that subjects tried to retrieve aspects of the story and then looked for the words to put in the appropriate slots of the retrieved aspects.

Meyer found that the frequency of recall of an idea unit high in the paragraph structure was determined by the unit's functional relationships to other units, rather than its content per se. This was not true, however, for idea units that were low in the content structure. In other words, two different passages with the same structure of specific relations, but with different content, would yield a high positive correlation of probability of an item's recall from the same point in the structure, so long as that point is a "high" point; items from the same point low in the structure are not correlated with respect to recall. In these cases, the words recalled are quite idiosyncratic to story and subject.

From this result and ones mentioned earlier, Meyer concludes that elements "high" in the passage are given more processing time than elements lower in the passage. Perhaps all higher-level material is attended to and with the remaining "processing-time" the subject attends to that low-level information which "catches his eye", (i.e., those units that seem inherently interesting). This would explain why, at low levels, structure does not predict recall. The mechanism Meyer suggests to explain differential retention is additional rehearsal. It has been shown, though, (Bjork & Jongeward, 1973) that additional rehearsal, per se, does not improve recall. One possible process might involve generating inferences and elaborations which draw upon previous knowledge of related information; the information higher in the structure would be elaborated more often than low-level information.

There is another explanation for her differences in recall which does not posit differential processing at input. Low-level information may have been stored equally well, but not recalled as well due to *output* interference (Roediger, 1973, 1974; see Crowder, 1976, for a discussion).

Meyer's representations do not include any inferences or elaborations, not even ones necessary to connect various propositions of different sentences. Clearly it is not her goal to represent the structure of prose in the reader's head; if that were her goal, she would not include every detail from the passage and she would include many inferences and details people integrate into prose when they read it. Since this is not the intent of her representation, one might ask what her structure is to represent. The fact that the structure looks like an outline strongly suggests that she represents what the author wanted to say, not what the reader absorbs. Her representation can be used to predict recall and as a metric to compare passages. While these aspects are useful, representing the author's intended message is less interesting to psychologists than is representing the reader's comprehension of it. Meyer's representation might have been more interesting if she had indicated how the author's intended meaning is transformed into prose and then what portion of the original meaning is represented in the "comprehender's" long-term memory. This might have predicted recall still better.

Meyer's thesis has been a major contribution to the study of prose processing and retention. She has developed the first representation that actually seems concerned with passage level information rather than sentence level material. Her representation has been empirically validated and will be a useful tool in the design of future recall experiments using prose. For example, it is possible that various manipulations, such as priming questions, which affect a subject's ability to answer questions, will interact with the information's "height" in the passage representation. Presumably an important point will be less affected by a focusing manipulation than would a detail.

Formal Models of Text Structure

There have been a number of attempts to identify the general organizational structures in prose (e.g., Colby, 1973; Lakoff, 1972; Mandler & Johnson, 1977; Propp, 1958; Rumelhart, 1975; Thorndyke, 1977; van Dijk, 1977). Developing a grammar for discourse has some of the same attractions that work on a grammar for sentences has, viz., the task is well defined, and it is therefore quite easy to determine when passages fail to fit the grammar.

Theory of Discourse and Macro-structures

Van Dijk (1975, 1977; Kintsch & van Dijk, 1975; van Dijk & Kintsch, 1976) has been concerned with a general theory of discourse. Van Dijk has incorporated Kintsch's work discussed above (Kintsch, 1974) into his theory of semantic representations for sentences and sequences of sentences. This representation is called the micro-structure of the passage. Van Dijk has also developed a theory of semantic representation for global discourse structures. The global structures are called the macro-structures of the passage. Macro-rules allow one to relate the micro-structures to the macro-structures. Macro-structures are the main aspects of a text that are stored during comprehension.

One reason van Dijk devised macro-rules and investigated macro-structures may have been due to his experimental results. He asked subjects to read a story and later recall it. Other subjects were asked to read and summarize the same passages. He found a great deal of consistency in the recalls and in the summaries. The common propositions found in all subjects' recall formed a sort of story itself and it was quite close to the summaries from the other subjects. These consistently recalled propositions are considered to comprise the macro-structure. Propositions most likely to be recalled either introduce main characters, give the characters goals, or describe the actions leading to these goals and results from the attempt. Propositions concerned with settings, mental actions, etc. tend to be ignored.

In van Dijk's analysis, the micro-structure usually entails the macro-structure. He derives three macro-rules: generalization, deletion and construction. These are posited to transform a text base into the core or essential macro-text base. The rule of *generalization* allows one to summarize a sequence of statements. There exist different levels of generality of abstraction entailed by the level of explicit text. The level of abstraction that is taken is the least general which also defines the smallest superset. The abstraction process stops when a superclass would not shorten the number of propositions. For example, statements in a text such as "Peter has a dog, and three cats. He has a parakeet, etc." would be reduced to "Peter has pets," not "Peter has animals" or "Peter has things".

Inferences using the *rule of construction,* on the other hand, are probably true but not necessarily true. Construction refers to finding a summary term for a set of actions. With this rule, the summary term is not *entailed* by the actions, but if all the required actions are stated, one infers the event. For example, "Peter laid foundations, built walls, built a roof . . ." becomes "Peter built a house." The other important macro-rule, the *rule of deletion* does not involve inferences. It follows the principle of irrelevancy, i.e., delete propositions that have no consequence, that are merely details. For example, one may delete from "Mary hit a blue ball that broke a window", the detail that the ball was blue.

Van Dijk and Kintsch have continued to explore the usefulness of the notion of macro-structures (van Dijk & Kintsch, 1976; Kintsch & van Dijk, 1975). They are not interested in the role of inferences, per se, but wish to demonstrate the importance of macro-structures to comprehension. However, their results also demonstrate how the reader uses his conceptual structures and world knowledge to infer what must be the appropriate organization of the text. Subjects read passages with the intra-paragraph organization intact, but the inter-paragraph organization scrambled for half of the readers. They felt that this essentially scrambled macro-structures without disturbing micro-structures. Although subjects took much longer to read stories with distorted macro-structures, the time to generate summaries was not significantly longer for the scrambled group. The summaries made by subjects who read scrambled passages could not be distinguished (by other subjects) from summaries of subjects who read normal passages. These results are consistent with the notion that subjects work harder to assemble or infer the macro-structure when the paragraphs have been scrambled (hence, the longer reading time), but that once this is done, the representation will be the same as that for an unscrambled passage. This study resembles that of Kintsch and Monk described earlier.

In a related experiment with children, Kintsch has found a great deal more consistency in recall of the important (macro-structure) propositions of a story when the story was presented with the macro-structure intact. Both this study and the one just described above, were taken as evidence of the existence of macro-structures. Although the notion of macro-structures is a useful one and the experiments are suggestive, one should ask how many of the possible set of outcomes from their experiments would be considered as consistent with the notion of "macro-structures". Note that in the case of the children's data, a failure to make use of the scrambled order was considered as evidence for the existence of macro-structures in children; for adults, the ability to overcome the scrambled order and produce a summary indistinguishable from those of normal passages is considered as evidence for macro-structures, too.

In a more recent paper, Kintsch and van Dijk (1978) have further articulated a model of text comprehension. This model is more of a processing model than their earlier work. The text base is postulated to be constructed "by a process operating in cycles and constrained by limitations of working memory." The model specified three types of operations: the organization of the meaning elements of a text into a coherent whole; the condensation of full meaning into gist; and the generation of new texts from what was retained through comprehension. The nature of the "semantic structures," viz., macro- and micro- structures, critical to the model is carried over from earlier work. Referential coherence is still important to the derivation of the structure of the text base. The application of macro-rules are now thought to be under the control of schemas.

Kintsch and Vipond (1978) have produced a way of determining readability of a passage that takes into account characteristics of the reader. In other words, the ordering of difficulty of passages can vary across readers depending on the capacity of a reader's short-term memory and size of the chunks in STM.

Story Grammars

Rumelhart. Rumelhart's (1975) generative grammar for stories and set of summarization rules was one of the first text grammars to be developed and tested and has influenced the work of other psychologists (e.g., Mandler & Johnson, 1977; Stein & Glenn, 1977; Thorndyke, 1977). In his chapter in Bobrow and Collins (1975) Rumelhart listed two sets of rewrite rules, one called syntactic rules, the other semantic interpretation rules. The important differences between the two sets of rules is not obvious. Syntactic rules specify how a sentence is decomposed; the semantic rules are intended to specify the relations among the parts.

Essentially a story is thought to consist of an episode in a particular setting. An episode consists of an event plus a reaction, where an event might consist of a change of state or an action, etc. Rumelhart gave several examples of stories and fables that seem to fit his grammar. He does not claim that the grammar will fit all stories, especially "more complex multi-protagonist stories". However, he states that it will account for a wide range of simple stories.

I had only limited success trying to apply Rumelhart's grammar to passages that I consider simple stories. For example, grammar allows for more than one episode; however, a story is allowed only one setting. Many simple stories have, in fact, different settings for different episodes. More important, the tree structure that is generated from the grammar fails to capture the relative importance of various episodes. That is, a second episode must be subordinate to the first. In some stories, however, a second episode can seem at least as important as the first even if the first episode caused the second.

The representation Rumelhart offers fails to capture critical inferences or interpretations that readers derive from a story. For example, in one fable that he represents (Rumelhart, 1977), a dog loses his meat through greed. The dog saw his reflection in the water, and thinking it was a different dog with meat, he snapped at the second piece of meat, only to lose his own. The reader comprehends the passage by noting something like "greed is counter-productive" or "be satisfied with what you have". Rumelhart's schema does not account for this derived, higher-level notion.

Recall data and summarization protocols have been examined for sets of stories that fit Rumelhart's grammar. Information high in his tree structures should be recalled more often than information (details) in the lower branches. Rumelhart considers the data of Thorndyke (1977) and Meyer (1975) as support for this basic idea. He has also collected data of his own looking at subjects' recall to stories after the subjects have summarized the stories. He found that recall was highly correlated with level in the representational hierarchy.

Although Rumelhart has accounted for much of the variance in subjects' recall using his grammar, one must still be cautious in making conclusions on the basis of its predictive power. All the stories he used in his experiment he designed (or selected) to fit the grammar (it is easy to find stories that will not fit the grammar). That is, before giving the stories to subjects to summarize and recall, he made sure that his own summary of the story fit the grammar.

Thorndyke. Thorndyke's grammar is adapted from Rumelhart's and bears a great deal of resemblance to it, except he emphasizes the importance of goals more than Rumelhart. The main character is seen as trying to solve a problem or obtain a goal. He demonstrated the importance of goal structure by rewriting stories that initially fit the grammar to violate various aspects of it. Recall was best in the untouched version of a story for which the goal structure was clearest. It was somewhat worse if the theme or goal was transposed to the end with all the subgoals removed; it was worse yet with no theme; the version that changed the story to consist only of stative and single action sentences had the poorest performance.

Mandler and Johnson. Mandler and Johnson (1977) have adapted Rumelhart's grammar and have constructed a story grammar (or "schema" as they call it) which seems to have a lot of promise. Their grammar allows more recursion of episodes and is somewhat more complex. They have deleted Rumelhart's semantic structure rules and added a moral at the topic level of the tree structure, similar to van Dijk's analysis.

Mandler and Johnson view the story schema as a collection of expectations that are modified as new semantic and syntactic information is processed. The schema is thought to facilitate encoding and retrieval. Because their model attempts to utilize general knowledge as well as knowledge of story grammars, it may be more successful in identifying the constituents of a story. Thorndyke seems to focus on plot structure almost to the exclusion of story content, while Mandler and Johnson analyze which aspects of an episode are most important to the reader. Their predictions concerning recall take content into account.

The experimental predictions of Mandler and Johnson go beyond the traditional ones that are connected with level in the hierarchical representation. For example, they postulate that causally connected episodes will be better recalled than temporally connected episodes. Their model includes several deletion transformations. The structure of the famous passage "The War of the Ghosts" used by Bartlett (1932) is an example of a structure that is far from ideal. Mandler and Johnson predicted which aspects of this passage would be forgotten and how modifications of the passage will affect recall. Their predictions concerning the types of information that will be retained was supported by a developmental study. The same trends were observed at all ages, although there were some qualitative differences, e.g., young children almost never recall reactions to events.

Stein and Glenn. Stein and Glenn's approach bears a great resemblance to that of Mandler and Johnson. It shares many of its virtues. Both grammars contain rules that define the units in a story and the relations among the units. Unlike Rumelhart or Thorndyke, they allow actions as well as states in their settings. One critical difference noted by Black (1978) is that the Stein and Glenn grammar does not allow lower level units to be re-written into higher level units, while the others do. Like Mandler & Johnson, they have validated their grammars with developmental data as well as adult data.

An example of the converging experiments Stein and Glenn have conducted to support their grammar are a series recently published by Glenn (1978). In some studies, she would vary the length of stories, keeping structure constant; in others, she would vary the structure of episodes. Amount recalled and number of elaborations generated was affected by length, but recall organization was not. Output protocols were scored according to which aspect of an episode was recalled more often than another. A consistent pattern emerged such that certain informational categories (e.g., internal response of protagonist vs. reaction of protagonist to a consequence) were recalled more frequently than others.

Stein and Glenn claim that their grammar is "inherently" a theory of processing. Glenn has gone on to argue that because the pattern of recall for a story is consistent with the story grammar's representation of it, "the story grammar, as a model of processing", is supported. Neither assertion is true. A grammar does not implicitly contain a process model. A processing model is necessary to predict behavior using a grammar. A story grammar only offers rules for generating acceptable stories, provided that one knows what the components in the grammar refer to.

Discussion of story grammars

The work on story grammars focuses on syntactic structures common to some simple stories. All of the theorists concerned with story grammars and structural representations of text make cogent arguments concerning the importance of such analyses. Mandler and Johnson state that their long range goal is to be able to predict what people will and will not be able to remember from connected discourse. It is also claimed that story grammars are needed to explain prose comprehension. However, an adequate process model has not been offered to explain how these grammars are involved in comprehension. For example, if the grammars proposed are actually used to help parse a passage, how does the reader know which of the many grammars (fables, narratives, complex stories, texts, etc.) is appropriate for a given output? How does the reader know when a passage is in violation of a given structure (i.e., unacceptable) as opposed to being an exemplar of another grammar? These are certainly not insurmountable problems, but they have not been touched by those arguing for the usefulness of these grammars.

Schank (1972) argued that syntactic parsers were not adequate to account for humans' ability to comprehend sentences. He argued that parsers should dwell on the meanings of words, not the form classes of the words. Similarly, comprehension models for passages should dwell on the meaning of sentences and not the possible structural representation. Mandler and Johnson admit that thus far they have stressed syntactic aspects of stories, largely ignoring the semantic content. They have incorporated the "content" of a passage into their theory more than most others, but they need to do much more.

The representation of a story in the comprehender's mind must result from some process operating upon the input. Some of the representations discussed thus far (e.g., Meyer) are intended to reflect the input as the *author* intended to convey the message. Other grammars discussed are intended to be predictive of the representation of the input in the comprehender's memory. Exactly how the grammar is represented in the comprehender's memory is not clear, nor is it clear how the grammar is used in building a representation.

USE OF WORLD KNOWLEDGE TO AID COMPREHENSION

Schemata, Frames and Prototypes

The introduction of the term "schema" to replace the term "grammar" is not unique to Mandler and Johnson. Rumelhart, (1976) and Rumelhart and Ortony (1977), have more recently viewed Rumelhart's general plot structure or story grammar as just one grammar or "schema" among many that people have available for everyday functioning. According to Rumelhart and Ortony there are schemata for pattern perception (e.g., a face schema), for motor skills (e.g., a juggling schema) as well as for comprehension (e.g., a "give" schema).

The schemata for stories are like a grammar in that they allow one to decide what is and what is not a story and what are the constituents of the story. If a story does not fit any of the available schemata, it is not an "acceptable" story. The schema Rumelhart describes is one in which something happens to the protagonist which sets up a goal. The rest of the story is concerned with trying to achieve the goal. A passage is a story fragment when a goal has been set up and the passage does not have a resoluton or an outcome.

The idea of using schemata is actually an old one. Rumelhart and Ortony, mention that Kant (1781) as well as Bartlett (1932) put forth related views. More recently, Chafe (1976) and Winograd (1977) have also adopted the term in trying to describe the mechanisms involved in prose comprehension. According to these theorists, a schema or prototype gives the basic structure of a class of its instantiations. No instantiation matches the prototype or schema perfectly. Chafe believes that when a person stores information about an event in memory, the person encodes it with respect to his prototype of similar events, but that the match is not all or none. He argues that information is stored in a "schematic" rather than a "propositional" form because a person does not report the same experience in the same way on two separate occasions, each report being a variation on the schema. This argument is less than convincing for two reasons. First, it is not clear why a propositional representation is incompatible with novelty in expressing the same event; second, my own impression is that the overlap in the retelling of events, jokes, etc. by a given individual is often more striking than any novelty in his exposition. In fact, such retellings often remind me of a prerecorded message being played again.

Winograd (1977) has also adopted the term schemata but relates it to Minsky's (1975) "frames". The distinction he offers between the two concepts is quite subtle. Frames are definitions; they denote a representation with variables or slots to be filled by a particular instance. Certain slots are optional, i.e., they do not have to be given a value for a particular instance. Others are mandatory and assume default values when not specified in the story. Schemata, on the other hand, are more flexible in that they allow for partial fits in expressing ideas.

Like Rumelhart, Winograd believes that schemata represent concepts which vary in their levels of abstraction; one schema can be embedded within another. Both Winograd and Rumelhart mentioned briefly when presenting papers at the 1976 Carnegie Symposium that schemata include procedural information for recognizing whether something should be classified as an instance of the schema. How this recognition process would work is not yet clear. (Bobrow and Winograd (1977) have developed a language called KRL (Knowledge Representational Language) that is supposed to recognize instances of a schema. The details of how this is accomplished have not been given, nor are limitations of the program known.)

Necessary Inferences

The notion that drawing inferences is necessary for comprehending a passage has been getting more attention in the last few years (e.g., Charniak, 1972; Clark, 1975; Frederiksen, 1977; Paris, 1975; Rieger, 1974; Schank & Abelson, 1975; Trabasso & Nicholas, 1977). Most of these papers demonstrate the importance of world knowledge and the drawing of inferences by presenting example stories and showing what inferences were necessary for comprehension. The papers by Clark, Rieger, and Trabasso and Nicholas actually present classifications of necessary inferences.

Clark. Clark (1975) and Clark and Haviland (1976) focus their analysis of inferences on those they call "authorized inferences." These inferences are called "authorized inferences" to indicate that they are intended by the speaker, or "backward inferences" to indicate that they are to connect the current sentence with previous ones. Many of the notions Clark and Haviland discuss are based on Grice's (1967) Co-operative Principle. One aspect of this principle concerns finding the intended antecedent that allows comprehension of the current statement. Finding the antecedent sometimes involves "bridging" or the drawing of inferences by the comprehender. Clark (1975) describes a number of types of bridging, some of which are described below.

One form of bridging is determining reference. Another type of bridge involves inferring information about something previously mentioned. This may also resolve reference. If one reads "I walked into the room. The ceiling was very high", it is easily inferred that "ceiling" refers to the ceiling of the room since it is a necessary part of the room. However, if one reads, "I walked into the room. The chandeliers sparkled brightly," the idea that the room has chandeliers has been induced; this time the bridge has conveyed new information. Chandeliers are not a standard attribute in a "room frame".

Sometimes the bridge, i.e., the number of inferences involved in making a connection, can be longer than in the above examples. From the set of sentences "John is a Republican. Mary is slightly daft, too," the implications would be: all Republicans are slightly daft; therefore, John is slightly daft. (Some people also infer that Mary is Republican, although that implication is less clear.) Those are both needed as antecedents to the "given information" that someone other than Mary is slightly daft. The number of inferences the listener assumes he should make to bridge the sentence is the *minimal number.* There are, undoubtedly, a large number of paths of inferences that could connect two sentences. The comprehender must use the rule of finding the shortest path; otherwise, comprehension would take a long time.

There are a number of experiments which support the notion that the time to comprehend a sentence is affected by the number of inferences that need to be drawn to do so, (Clark & Haviland, 1976; Clark, 1975). Some sentences that subjects read were preceded by a direct antecedent, others by an indirect antecedent. Comprehension of the identical sentence was faster when preceded by the direct antecedent. For example, subjects were faster to comprehend the second sentence in "Fran took the beer out of the car. The beer was warm," than in "Fran took the picnic supplies out of the car. The beer was warm," or in "Fran likes beer. The beer was warm," (the last being a control for the word "beer").

Clark also makes the distinction between *backward* and *forward* inferences and *authorized* and *unauthorized* inferences. Forward inferences are unauthorized (not demanded by the speaker) and are indeterminate in number (i.e., there is no minimal number rule). These are elaborations or embellishments of the input in the text. Backward inferences that are authorized and follow the minimal number rule are the bridges described above. Unauthorized, backward inferences follow from statements such as Nixon's "I am not a crook." In this case, the speaker did not intend the implication that there is reason to consider that he might be a crook.

Schank's model for drawing inferences

Clark writes as if there are very few and often only one backward inference to be made per sentence. However, Schank's model for sentence comprehension and for paragraph comprehension (Schank, 1972, 1973, 1975a), suggests that many such inferences are made. These backward inferences tend to concern presupposition and enabling conditions. The fact that a large number of inferences need to be made has always seemed to be one of the weakest aspects of these models. Schank (1972, 1973) asserted in his papers on conceptual dependence (sentence parsing) that during comprehension, the verb is decomposed into its semantic primitives. Decomposition includes inferring the instrument of action when the verb is one that "takes" an instrument. Even in sentences where the instrument is explicitly stated, more implicit instruments may need to be inferred, recursively. For example, the conceptual dependency representation of the instrument in "Fred hit Bill with a stick" would be that "Fred did something with a stick" and the instrument for that would be that "Fred grabbed the stick" which implies that "Fred moved his hand toward the stick" and so on. (At a Sloan sponsored workshop at Yale in June, 1978, Schank indicated that his position on decomposition has been modified. He now believes there exist "primitives" at many different levels. One need not unpack a verb down to the original "primitive acts" postulated by Schank. I find the notion of different levels of primitives somewhat disconcerting. Perhaps primitives now mean "concepts".)

In a similar vein, Reiger (1974), a student of Schank, developed a scheme in which a seemingly limitless number of "real world inferences" would be made during comprehension. As he states:

> "When the current (Conceptual Memory) is turned loose, it will often generate upwards of 100 inferences from a fairly banal stimulus such as "John gave Mary the book." (p. 31)

One motivation for Schank's decision to develop representations for paragraphs was that sentence comprehension involved the generation of too many inferences; embedding a sentence in a paragraph limits the number of relevant inferences. He feels that in paragraph comprehension, one only needs to generate those inferences that will connect sentences of the story together; one must find enabling conditions, i.e., causal chains that allow an event to occur. He distinguishes between absolutely and reasonably necessary conditions (ANCs and RNCs, respectively). Examples of ANCs for the phrase "John began to mow his lawn" would be John having a lawn, possessing a lawn mower, John being alive, etc. Reasonably necessary conditions would be that it is not cold, nor rainy, nor is there any snow on the ground, etc. None of these conditions need to be stated in the paragraph that includes the phrase. All these statements could be satisfied by "normality assumptions." If the ANCs can be satisfied by "normality assumptions", one only checks to be sure that none of these assumptions are violated by some other assertion in the paragraph.

Script, Plans and Goals

Schank and Abelson (1975, 1977; Abelson, 1975) proposed that people utilize situation-specific knowledge, called scripts, in order to make inferences and determine if something seems plausible or reasonable in a particular context. Normality assumptions will vary from context to context (or script to script). The model that uses scripts is implemented in a computer program that summarizes, paraphrases, and translates stories (Schank, 1975b). It can also answer questions about the story concerning what happened, when did something happen and why did it happen.

In addition to scripts, which capture the essence of a stereotyped sequence of actions or events for a well-known situation, there are other comprehension mechanisms such as plans. Plans handle novel situations. Plans describe the set of choices associated with accomplishing a goal. When a plan is frequently implemented it becomes a script. Most people do not have a "how to become president" script or a "what to do when the house burns down" script but could implement a plan. An example of a script common to most Americans is a restaurant script. Consider (1) below:

> (1) John went into a restaurant. He ordered a hamburger and a coke. He asked the waitress for the check and left.

Many aspects of this standard script have been omitted, but are automatically inferred. If the reader of (1) were asked at a delay "Did John eat a hamburger?", the response would probably be yes although that was never explicitly stated.

Scripts are a set of scenes and within each scene there is a causal chain such that each action *enables* the next. For example, in the restaurant script, asking for the check *enables* the patron to receive the check. Many parts of a script can be modified and, in fact, frequently are. There are only a few aspects of a script, e.g., ingesting food, which are considered critical. Often there are deviations. Consider scenario (2):

(2) John went to a restaurant. He ordered a hamburger. It was cold when the waitress brought it. He left her a very small tip.

Since there are two deviations from the typical script, viz., "cold" and "very small", the processor would hypothesize that the deviations are related.

Even though we do not have scripts for every situation encountered in a story, we do need relevant conceptual structure to understand a given input and infer the appropriate connections. Consider the following passage:

(3) John knew his wife's operation would be very expensive. There was always Uncle Harry . . He reached for the suburban phone book.

It is unlikely that we have a "paying for an operation" script, yet we might have a "raising money for important expenditures" script. That is, the comprehension of (3) in terms of inferences made would not be very different if "son's education" were substituted for "wife's operation". In each case there is a general goal state. Trying to achieve a goal involves a plan.

A plan is a series of actions that will hopefully realize a goal. Much like the General Problem Solver (Newell & Simon, 1972), plans to achieve the goal state go through intermediate states, trying to reduce differences between the goal state and current state. Plans consist of a set of "deltacts", actions or subplans. When a set of actions is used often enough together, the set becomes a script.

Sometimes in the middle of a story, the actions that a character is pursuing will be frustrated and there will be an abrupt shift in plans. In order to better understand why the actor would shift from the one activity to another, Schank and Abelson described a theory of goal substitution, goal forms, goal initiations, etc. When the system cannot find an appropriate script, it tries to induce what the character's plans would be using the goal monitor. In addition to keeping track of the motivating influence of the character's goals, the monitor must recognize when a goal has been triggered, makes predictions about what events will be caused by it and keeps track of a goal's fate.

In Schank and Abelson's system, the representation of a story has three levels of description: (1) the lowest level is conceptual dependency (CD), described earlier; (2) main conceptualization, (MAINCONS), describe the important aspects of a scene end of a script; (3) the highest level is the knowledge structure (KS) which has the script, plan or theme information. The CD level is in some ways similar to Kintsch and van Dijk's micro-structure, while the MAINCONS are like the macro-structure. MAINCONS generate expectations. The three levels are connected by pointers. Like Kintsch and van Dijk and others, Schank and Abelson do not think the lowest level is involved in summarizing. Schank has also predicted that items most likely to be forgotten are those that are at "dead ends" in causal chains. However, unlike Kintsch and van Dijk, Schank has never empirically tested his ideas.

A major weakness of script theory is devising a reasonable scheme for evoking a script—how does the program know if a script is the correct one? Sometimes several scripts seem appropriate concurrently, (e.g., a restaurant script and a romance script). There is no mechanism that allows both scripts to be operating and to make predictions at the same time, let alone interact. These difficulties are not unique to script theory. All extant theories that rely on "bundles" of knowledge to help direct inferences and expectations, (e.g., frames and schemata), share the same dilemma. Scripts are currently instantiated when a key phrase has been mentioned (e.g., John *went to a restaurant*) and a line fitting the script has been mentioned. Most of the script recognition task has been delegated to the scripts themselves. Others working on simulations have also alluded to procedures attached to schemata that are evoked in the appropriate context so that the relevant schema or script is activated. The workability of these notions remains to be tested in implementation.

The Elaboration Theory

The script model described above was described in most detail because it is the processing model with the most promise, in my opinion. The script-elaboration model, a hybrid, borrows heavily from Schank and Abelson's work as the name implies. This hybrid is essentially their model with the addition of the notion of the importance of elaborative processing. The added assumption that the reader makes many embellishments during reading has a great deal of psychological plausibility and accounts for a large body of data that would otherwise be difficult to explain.

Elaborations are generated that are consistent with the script invoked, but not necessarily true. They tend to be highly idiosyncratic. Consider the following sub-passage.

(4) Bill took Joanne to a steak and ale place, not very fancy but all he could afford. Joanne tried to study the menu intently while Bill gazed at her.

References to "a steak and ale place" and "menu" would call up a restaurant script and various inferences would be added to the story structure such as Bill and Joanne were shown to a table, given menus by the hostess or waitress, etc. The reader might also elaborate that the restaurant was semi-dark, that they were seated in a booth, that the table had placemats rather than a tablecloth, that Bill was infatuated with Joanne, that Joanne was somewhat embarrassed by Bill's infatuation, that this story was a romance story, etc. Another reader might have another set of elaborations entirely, differing in the nature of what was embellished or how it was embellished.

Elaborations occur at many levels. They can be embellishments of physical descriptions or characters' personalities, imputing goals or intentions to the author, etc. Low level elaborations are generated by accessing from long-term memory salient exemplars of concepts referenced in the story. That is, as each word in a story is read, the corresponding concept in memory may be activated which, in turn, evokes associations to the concept. The instance of a concept stored in memory that would be used to embellish the story will be the most common instance that is consistent with the context described in the story.

The generation of elaborations is similar to processes used when scripts are employed to aid comprehension. In both cases, past experience is used to understand a current situation and in both cases, more information is added to the memory representation of the story. The difference between inferences that represent the omitted main conceptualizations of a script and elaborations is two-fold: Any omitted script action is likely to be inferred and the default value for the omission, (i.e., the inference) is likely to be similar across readers. On the other hand, the number of elaborations generated varies greatly with circumstance and the content of elaborations varies across readers. How much one elaborates depends upon previous experience with the material, inherent interest in the subject matter, understanding of the text, time allotted to read it, concentration and general tendency to elaborate.

Since elaborations are only moderately plausible and are optional, one might ask why people generate elaborations and since they are not necessary, why one should include them in a comprehension model. They serve a number of basic functions, namely, finding connections among sentences, generating expectations about subsequent input, detecting anomalies and aiding retention.

Elaborations may help to detect anomalies. Assertions within a passage may not be in direct contradiction with one another, but the elaborations evoked by the sentences may be inconsistent. Consider (5):

(5) Alice went to Jimmy's house for lunch. Jimmy's mother served them tunafish sandwiches. Alice liked her sandwich very much and had almost finished it when all of a sudden, her dentures fell out of her mouth.

Although (5) seems anomalous, there is nothing in the passage that is inherently contradictory. Elaborations generated to it, however, might be in direct conflict. For instance, part of my elaborations involved Alice having the smooth skin of a young girl. Then when I imagined dentures falling out, I elaborated a wrinkled face with an old mouth and exposed gums. The anomaly was detected by noting the contradiction of these two facial features. There are of course many other sets of inferences that would also contradict.

Perhaps the most important aspect of the process of elaboration is the benefit to long term retention of the input. Any particular proposition is fragile and may become irretrievable at test. However, if an input proposition had been richly elaborated during reading, then only a few of the redundant propositions need be retrieved at recall to infer the gist of the original input. Suppose, for example, that a reader elaborates passage (4) as follows: "I imagine that the steak and ale place looks a little like Stickney's Restaurant. Perhaps Joanne only pretends to be ingrossed by the menu because she is embarrassed by Bill's attentiveness . . . " etc. Even if the reader could not recall any of the input propositions, he would still have a good chance of recalling the gist of the passage from recalled elaborative fragments. This reconstructive view of memory is not unlike that advocated by Bartlett (1932).

Experimental Evidence in Support of the Elaboration Theory

From the point of view of providing strong empirical support for the elaboration theory, a serious problem arises in that the theory implies that an experimenter has poor ability to manipulate the amount and direction of elaboration. Nonetheless, there are a number of experiments concerning selectivity in memory for prose which are consistent with an elaboration-plus-reconstruction viewpoint. If subjects have more ability to make certain types of elaborations than others or if subjects are directed to make certain elaborations rather than others, one should see better memory for material in the direction of the preferred elaboration and more distortion of material in the direction of the preferred elaboration. In the classic study conducted by Bartlett (1932), subjects from pre-World War I England studied a northwest coast Indian story, "The War of the Ghosts." Bartlett obtained what he interpreted as systematic distortion of the material in the direction of the knowledge of his subjects. This distortion took the form of additions to the material that made the story more consistent with the world view of his subjects, deletion of inconsistent information, and transformations of inconsistent information to make it more consistent with prevailing beliefs.

There has been a long history of debates (e.g., Gould & Stephenson, 1967; Anderson & Bower, 1973; Spiro, 1975) over the extent to which Bartlett's subjects were really misremembering and the degree to which they were knowingly confabulating in response to perceived task demands. It seems that, at least to some degree, subjects are aware of their distortions and are able to assign lower confidence to these than to veridical recalls. However, this debate misses an important point: The behavior of subjects in Bartlett's task is typical of prose processing. Normally, the reader does not make distinctions between what was actually read in a passage and what is a plausible inference. With most stories the inferences made are plausible extensions of the story and are not distortions. It was Bartlett's clever story selection that served to highlight the elaborative behavior of subjects. The experiments described earlier by Sulin and Dooling (1974) and Dooling and Christiaansen (1977) that showed greater confusion to thematically related foils, suggests that distortion can occur as a reconstructive process of an encoding process.

Bower (1976) reports an interesting experiment looking at the effect of prior information on subjects' memory for a passage. Subjects were given a story that consisted of episodes. Half of the subjects were given prior information that would suggest an unusual interpretation of some of the subpassages. For instance, subjects might be told that the main character (a college co-ed) had just found out that she was pregnant. The story follows the heroine through five episodes: making a cup of coffee in the morning, visiting a doctor, attending a lecture, going shopping in a grocery store, and attending a cocktail party. The meaning of the episodes can be very different depending on whether or not we view the heroine as pregnant.

Subjects given the interesting interpretation recalled many more inferences appropriate to the pregnancy theme. However, they also recalled more of those episodes related to the theme. This result is what would be expected if subjects had used the information about pregnancy to elaborate. These elaborations should make the text information more redundant and introduce additional inferences.

Hayes (Personal communication, July, 1977) has found a similar correlation between number of intruded inferences and overall memory for text. Hayes and his colleagues tried to find out what mechanisms allow some people to remember more than others. They pre-tested subjects on their memory for various historical facts and then classified them as those who remember a lot of history and those who do not. The subjects were then given a fictitious history passage to read. The same subjects who knew more veridical history performed better on a test of the fantasy history passage. Subjects were also asked to free recall the passage that they had read. Not only did the subjects with better history memory recall more, they also "recalled" many elaborations that were not asserted. These elaborations were not simple paraphrases of the passage nor were they simple inferences. The subjects classified as having poor memory for history offered almost no elaborations. From this finding, Hayes conjectured that embellishing the input with elaborations promotes better retention.

Brown, Smiley, Day, Townsend, and Lawton (1977) conducted a study with children that supports the elaboration theory, by showing that the types of elaborations generated can be manipulated and that generating more elaborations improves recall. The study indicates that if teachers provide students with background knowledge, students are likely to remember more of the material presented to them. In the Brown, et al. study, children in various grades were presented with information about a fictitious tribe called the "Targa" or learned about people from Spain. Those who learned about the Targa, were either told that the tribe consisted of Eskimos or desert Indians. A week later all groups read a story about a young boy from the Targa tribe, and no mention was made about what they had studied the preceding week. Of those receiving relevant background information, intrusions and biases in interpretations of ambiguous sentences were consistent with the orientation given earlier. More important, those subjects in the Spanish control recalled significantly less of the veridical material. The usefulness of the background material was evident at all age levels. However, not only did older children recall significantly more veridical information, they recalled significantly more elaborations (had more intrusions) consistent with the background material. In other words, both recall and number of elaborations increased with age.

In a study of mine, (Reder, 1976, 1979), I manipulated more directly the amount of elaboration given to prose material. An earlier study of mine indicated that subjects have very good memory for sentences used in the stories used in the experiments. So, the dependent measure chosen was the *speed* with which subjects can make judgments about a story, not the accuracy. The task demanded of subjects was to make plausibility judgments, which seems to resemble everyday tasks more than verbatim memory judgments. Three factors were manipulated orthogonally. With the use of the additive factors logic, I determined that subjects do not search for a specific fact, but rather for a set of propositions that will be relevant for judging plausibility. The speed with which these propositions become available is a function of the number of elaborations generated.

Just as the number of relevant elaborations is postulated to affect *retention* of the input, amount of relevant elaborations is thought to affect the *speed* of retrieval of relevant information. This notion is based on the assumption that searching through memory for relevant information takes time and that the greater the proportion of relevant information to irrelevant, the faster a useful or relevant fact can be found; hence, faster RTs occur with more elaborations. (See Reder, 1976, for a fuller discussion.)

Discussion

The approaches represented in this last section tend to differ from others in one of a number of respects. The artificial intelligence models are either implemented on a computer or are purported to be sufficiently specified to be implemented. Most of the models in this section are more concerned with the reader's use of world knowledge than the use of knowledge of language to understand a passage.

These qualities make the models in this last section important advances. When models are implemented, inherent weaknesses in them and issues that were overlooked are quickly brought into focus: A bad program will not run or will produce the "wrong" behavior. In other words, a program serves as a check on the viability of the model that generated it. On the other hand, sometimes the motivation for a particular feature of a model may be computational ease, and the feature may not seem psychologically plausible. Often when computer implementations of models do not work properly, patches are made to the program to make them work. When more and more patches are added to the program, it becomes incomprehensible as a model; the modifications to the program tend to have no psychological plausibility whatsoever. Furthermore, AI models are not always programmed and thus one has neither empirical support nor a computer implementation as confirmation of their viability.

Despite these criticisms, I feel the research in this section is the most commendable. It incorporates world knowledge and is concerned with processing the input not merely representing it. The most useful theories will be those that predict (1) what information will be best retained from a story, and (2) how to improve a reader's memory for and comprehension of a passage. Approaches concerned with story grammars seem less able to address (2).

The story grammar theories and the processing models share a common flaw: few seem to generalize to most texts or stories. If these theories are to have practical applications to education (or even be of more theoretical interest), they must strive to handle a larger class of prose. Hopefully, the complexity of a model will not keep pace with increases in generality.

SUMMING UP

This literature review has stressed the need for more attention to the elaborative aspects of comprehension. The inferences and elaborations that are made during comprehension should be reflected in a structure that is intended to represent how a passage is stored in memory, and a mechanism should be specified for how these inferences are made and why they are made. The hybrid model proposed is only a skeleton of a complete theory of prose comprehension. The exact structure of scripts, how scripts are related in memory, what other memory structures exist, how representations of stories are stored, etc. all need to be further specified.

Despite the sketchiness of the hybrid model, there are a number of conclusions that can be made that should be of interest to anyone involved in the theory of or application to prose comprehension. A person's comprehension and memory for a passage can be affected by a number of manipulations. The "inherent" differences among readers seems to reflect, in part, differential propensities to elaborate; nevertheless, educators should continue to try to influence the extent of elaboration during reading. Experiments have shown that certain questions asked in the appropriate position can influence the memory for certain facts. Prior knowledge, and inherent interest in the subject matter, orienting instructions, etc. can also affect retention and comprehension.

ACKNOWLEDGEMENTS

An early draft of this paper was prepared while the author was an NSF graduate fellow at the University of Michigan and later ones while an NIMH post-doctoral fellow at Yale University supported by a contract from NIE. The manuscript has benefited from the comments of J. Anderson, J. Baker, S. Chipman., N. Stein, and R. Tournageau. The first draft was written in 1975-76. The last "update" of research was 1977-78. A longer version of this review has appeared in the *Review of Educational Research*, Spring, 1980, *50*, 5-53. The author's address is Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

REFERENCES

Abelson, R. P. Concepts for representing mundane reality in plans. In D. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science.* New York: Academic Press, 1975.

Abelson, R. P. Script processing in attitude formation. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior.* Hillsdale, New Jersey: Erlbaum, 1976.

Anderson, J. R., & Bower, G. H. *Human associative memory.* Washington, D. C.: Hemisphere Press, 1973.

Anderson, J. R., & Reder, L. M. An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, N. J.: Erlbaum, 1978.

Anderson, R. C. How to construct achievement tests to assess comprehension. *Review of Educational Research,* 1972, *42,* 145-170.

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Anderson, R. C., & Carter, J. F. Retroactive inhibition of meaningfully-learned sentences. *American Educational Research Journal,* 1972, *9,* 443-448.

Anderson, R. C., & Pichert, J. W. *Recall of previously unrecallable information following a shift in perspective* (Tech. Rep. No. 41). Champaign-Urbana: Center for the Study of Reading, University of Illinois, 1977.

Ausubel, D. P. The use of advance organizers in learning and retention of meaningful verbal material. *Journal of Educational Psychology,* 1960, *51,* 267-272.

Ausubel, D. P., & Fitzgerald, D. Organizer, general background and antecedent learning variables in sequential verbal learning. *Journal of Educational Psychology,* 1962, *53,* 243-249.

Bartlett, F. C. *Remembering: A study in experimental and social psychology.* Cambridge: University Press, 1932.

Bjork, R. A., & Jongeward, R. H. *Rehearsal and mere rehearsal.* Unpublished manuscript, University of Michigan, 1973.

Black, J. *Theories of story memory structure.* Unpublished manuscript, Stanford University, 1978.

Bobrow, D. G., & Winograd, T. An overview of KRL, a knowledge representation language. *Cognitive Science,* 1977, *1,* 3-46.

Bobrow, S., & Bower, G. H. Comprehension and recall of sentences. *Journal of Experimental Psychology,* 1969, *80,* 455-461.

Bobrow, D. G., & Collins, A. (Eds.), *Representation and understanding: Studies in cognitive science.* New York: Academic Press, 1975.

Bower, G. H. Selective facilitation and interference in retention of prose. *Journal of Educational Psychology,* 1974, *66,* 1-8.

Bower, G. H. *Comprehending and recalling stories.* Section 3, Presidential Address, American Psychological Association, Washington, September, 1976.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology,* 1971, *2,* 331-350.

Bransford, J. D., & Johnson, M. K. Considerations of some problems of comprehension. In W. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973.

Bransford, J. D., & McCarrell, N. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, New Jersey: Erlbaum, 1975.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development*, 1977, *48*, 1454-1466.

Bruning, R. H. Effects of review and testlike events within the learning of prose material. *Journal of Educational Psychology*, 1968, *59*, 16-19.

Carpenter, P. A., & Just, M. A. Reading comprehension as eyes see it. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

Carroll, J. B. *Learning from verbal discourse in educational media: Review of the literature* (Final Report Project 7-1069). Princeton, New Jersey: Educational Testing Service, 1971.

Chafe, W. L. Creativity in verbalization and its implications for the nature of stored knowledge. In R. Freedle (Ed.), *Discourse production and comprehension*. Hillsdale, N.J.: Erlbaum, 1976.

Charniak, E. *Toward a model of children's story comprehension*. MIT Artificial Intelligence Laboratory Mimeo, December 1972.

Clark, H. H. Bridging. In R. Schank & B. Nash-Webber (Eds.), *Theoretical issues in natural language processing*. Proceedings of conference at the Massachusetts Institute of Technology, June 1975.

Clark, H. H., & Haviland, S. E. Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse production and comprehension*. Hillsdale, N.J.: Erlbaum, 1976.

Colby, B. A partial grammar of Eskimo folktales. *American Anthologies*, 1973, *75*, 645-662.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 240-247.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671-684.

Crothers, E. Memory structure and the recall of discourse. In R. Freedle & J. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D. C.: V. H. Winston, 1972.

Crowder, R. G. *Principles of learning and memory*. Hillsdale, N.J.: Erlbaum, 1976.

Dawes, R. Cognitive distortion. *Psychological Reports*, 1964, *14*, 443-459.

Dawes, R. Memory and the distortion of meaningful written material. *British Journal of Psychology*, 1966, *57*, 77-86.

Deese, J., & Kaufman, R. Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Educational Psychology*, 1957, *54*, 180-187.

Dooling, D. J., & Christiaansen, R. E. Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 428-436.

Dooling, D. J., & Lachman, R. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, *88*, 216-222.

Dooling, D. J., & Mullet, R. L. Focus on thematic effects in retention of prose. *Journal of Experimental Psychology*, 1973, *97*, 404-406.

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart and Winston, 1968.

Frase, L. T. Learning from prose material: Length of passage, knowledge of results and position of questions. *Journal of Educational Psychology*, 1967, *58*, 266-272.

Frase, L. T. Questions as aids to reading: Some research and a theory. *American Educational Research Journal*, 1968, *5*, 319-332.

Frase, L. T. Paragraph organization of written materials. The influence of conceptual clustering upon level of organization. *Journal of Educational Psychology*, 1969, *60*, 394-401.

Frase, L. T. Influence of sentence order and amount of higher level text processing upon reproductive and productive memory. *American Educational Research Journal*, 1970, *7*, 307-319.

Frase, L. T. Effect of incentive variables and type of adjunct questions. *Journal of Educational Psychology*, 1971, *62*, 371-375.

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In R. O. Freedle, & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D. C.: V. H. Winston, 1972.

Frase, L. T. Prose processing. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Frase, L. T., & Kammann, R. Effects of search criterion upon unanticipated free recall of categorically related words. *Memory and Cognition*, 1974, *2*, 181-184.

Frederiksen, C. Effects of task-induced cognitive operations on comprehension and memory processes. In R. O. Freedle, & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D. C.: V. H. Winston, 1972.

Frederiksen, C. Acquisition of semantic information from discourse: Effects of repeated exposures. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 158-169. (a)

Frederiksen, C. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, *7*, 139-166. (b)

Frederiksen, C. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458. (c)

Frederiksen, C. H. *Inference and the structure of children's discourse*. Paper presented at the Symposium on the Development of Discourse Processing Skills, Society for Research in Child Development meetings, New Orleans, 1977.

Gagne, R. M., & Wiegand, V. K. Effects of superordinate context on learning and retention of facts. *Journal of Educational Psychology*, 1970, *61*, 406-409.

Glenn, C. G. The role of episodic structure and of story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 229-247.

Gould, A., & Stephenson, G. M. Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology*, 1967, *58*, 39-49.

Greeno, J. G., & Noreen, D. L. Time to read semantically related sentences. *Memory and Cognition*, 1974, *2*, 117-120.

Grice, H. P. Logic and conversation. The William James Lectures, Harvard University, 1967.

Grimes, J. *The thread of discourse*. The Hague: Mouton, 1975.

Hyde, T. S., & Jenkins, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 471-480.

Just, M. A., & Carpenter, P. A. Inference processes during reading: Reflections from eye fixations. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, NJ: Erlbaum, 1978.

Kant, E. [Critique of pure reason] (N. K. Smith, trans.). London: Macmillan & Co., 1963. (Originally published, 1781; 2nd ed., 1787).

King, D. R. W., & Greeno, J. G. Invariance of inference time when information was presented in different linguistic formats. *Memory and Cognition*, 1974, *2*, 233-235.

Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W. On recalling stories. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

Kintsch, W., & Keenan, J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, 1973, *5*, 257-274.

Kintsch, W., McKoon, G., & Keenan, J. Memory for information inferred during reading. In W. Kintsch, *The representation of meaning in memory* (Chap. 8). Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W., & Monk, D. Storage of complex information in memory: Some implications of the speed with which inferences can be made. *Journal of Experimental Psychology*, 1972, *94*, 25-32.

Kintsch, W., & van Dijk, T. A. Recalling and summarizing stories. (Comment on se rappelle et on resume des histoires), *Languages*, 1975, *40*, 98-116.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, *85*, 363-394.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In Lars-Goran Niesson (Ed.), Hillsdale, N.J.: Erlbaum, 1978.

Kirscher, M. C. *The effects of presentation order and repetition on free recall of prose*. Unpublished masters thesis, Cornell University, 1971.

Lakoff, G. Linguistics and natural logic. *Synthese*, 1970, *22*, 151-271.

Lakoff, G. Structural complexity in fairy tales. *The Study of Man*, 1972, *1*, 128-150.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

McGaw, B., & Grotelueschen, A. Direction of the effect of questions in prose material. *Journal of Educational Psychology*, 1972, *63*, 586-588.

McKoon, G. Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 247-260.

Meyer, B. J. F. *The organization of prose and its effect on recall.* Amsterdam: North-Holland Publishers, 1975.

Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology,* 1973, *65,* 109-117.

Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill, 1975.

Montague, W. E., Adams, J. A., & Kiess, H. O. Forgetting and natural language mediation. *Journal of Experimental Psychology,* 1966, *72,* 829-833.

Montague, W., & Carter, J. Vividness and imagery in recalling connected discourse. *Journal of Educational Psychology,* 1973, *64,* 72-75.

Myers, J. L., Pezdek, K., & Coulson, D. Effect of prose organization upon free recall. *Journal of Educational Psychology,* 1973, *65,* 313-320.

Myrow, D. L., & Anderson, R. E. Retroactive inhibition of prose as a function of the type of test. *Journal of Educational Psychology,* 1972, *63,* 303-308.

Newell, A., & Simon, H. *Human problem solving.* Englewood Cliffs, N.J.: Prentice Hall, 1972.

Owens, J. E., & Bower, G. H. *Character point of view in text comprehension and memory.* Paper presented at meeting of the American Psychological Association, San Francisco, August 1977.

Paris, S. G. Integration and inference in children's comprehension and memory. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. R. Lindman, & D. B. Pisoni (Eds.), *Cognitive theory* (Vol. 1). Hillsdale, N.J.: Erlbaum, 1975.

Paris, S. G., & Lindauer, B. K. Constructive aspects of children's comprehension and memory. In R. V. Kail & J. W. Hagen (Eds.), *Perspective on the development of memory and cognition.* Hillsdale, N.J.: Erlbaum, in press.

Paul, I. H. Studies in remembering. *Psychological Issues,* 1959, 1 (2).(Monograph 2).

Perlmutter, J., & Royer, J. M. Organization of prose materials: Stimulus, storage and retrieval. *Canadian Journal of Psychology,* 1973, *27,* 200-209.

Postman, L. Extra-experimental interference and the retention of words. *Journal of Experimental Psychology,* 1961, *61,* 97-110.

Propp, V. Morphology of the folktale. *International Journal of American Linguistics,* 1958, *24* (Part III).

Prytulak, L. G. Natural language mediation. *Cognitive Psychology,* 1971, *2,* 1-56.

Reder, L. M. *The role of elaborations in the processing of prose.* Unpublished doctoral dissertation, 1976.

Reder, L. M. The role of elaborations in memory for prose. *Cognitive Psychology,* 1979, *11,* 221-234.

Reder, L. M., & Anderson, J. R. A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior,* 1980, *19,* 121-134.

Richardson, P., & Voss, J. F. Verbal context and recall of meaningful material. *Journal of Experimental Psychology,* 1960, *60,* 417-418.

Rieger, C. J. *Conceptual memory: A theory and computer program for processing the meaning content of natural language utterances.* Unpublished doctoral dissertation, Stanford University, 1974.

Roediger, H. L. III Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 664-657.

Roediger, H. L. III Inhibiting effects of recall. *Memory and Cognition,* 1974, *2,* 261-269.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), *Learning and the educational process.* Chicago: Rand McNally, 1965.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal,* 1966, *3,* 241-249.

Rothkopf, E. Z. Structural text features and the control of processes in learning from written materials. In R. O. Freedle & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D. C.: V. H. Winston, 1972.

Rothkopf, E. Z., & Bisbicos, E. Selective facilitative effects of interspersed questions on language from written material. *Journal of Educational Psychology,* 1967, *58,* 56-61.

Rothkopf, E. Z., & Kaplan, R. Exploration of the effect of density and specificity of instructional objectives on learning from text. *Journal of Educational Psychology,* 1972, *63,* 309-313.

Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding.* New York: Academic Press, 1975.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. La Berge & J. Samuels

(Eds.), *Basic processes in reading: Perception and comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

Rumelhart, D. E. *Comprehension and the structure of brief stories.* Paper presented at the Carnegie Symposium on Cognitive Processes in Comprehension, May 1976.

Rumelhart, D. E., Lindsay, P., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

Rumelhart, D. E., & Levin, J. A. A language comprehension system. In D. Norman, D. Rumelhart, & the LNR Research Group. *Explorations in cognition.* San Francisco: W. H. Freeman, 1975.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Schallert, D. L. Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior,* 1976, *15,* 621-632.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology,* 1972, *3,* 552-631.

Schank, R. C. Identification of conceptualizations underlying natural language. In Schank & Colby (Eds.), *Computer models of thought and language.* San Francisco: Freeman, 1973.

Schank, R. C. The structure of episodes in memory. In D. G. Bobrow & A. M. Collins (Eds.), *Representing and understanding.* New York: Academic Press, 1975. (a)

Schank, R. C. & the Yale A. I. Project. *SAM—A story understander* (Report No. 3). New Haven: Yale University, August 1975. (b)

Schank, R. C. The role of memory in language processing. In C. Cofer & R. Atkinson (Eds.), *The nature of human memory.* San Francisco: Freeman, 1975. (c)

Schank, R. C., & Abelson, R. P. Scripts, plans, and knowledge. Paper presented at the 4th International Joint Conference on Artificial Intelligence, Tbilisi, USSR, 1975.

Schank, R. C., & Abelson, R. P. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, N.J.: Erlbaum, 1977.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Educational Psychology,* 1977, *69* (4), 381-387.

Spiro, R. J. *Inferential reconstruction in memory for connected discourse* (Tech. Rep. No. 2). Urbana, Ill.: University of Illinois, Laboratory for Cognitive Studies in Education, 1975.

Sulin, R. A., & Dooling, D. J. Intrusions of a thematic idea in retention of prose. *Journal of Experimental Psychology,* 1974, *103,* 225-262.

Stein, N. L., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multidisciplinary perspectives in discourse comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

Thomas, E. L., & Robinson, H. A. *Improving reading in every class: A sourcebook for teachers.* Boston: Allyn and Bacon, 1972.

Thorndyke, P. *Cognitive structure in human story comprehension and memory.* Unpublished doctoral dissertation, Stanford University, 1975.

Thorndyke, P. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology,* 1977, *9,* 77-110.

Trabasso, T., & Nicholas, D. W. *Memory and inferences in the comprehension of narratives.* Paper presented at a conference entitled "Structure and process, models in the study of dimensionality of children's judgements," Kassel, Germany, June 1977.

Tulving, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review,* 1973, *80,* 352-373.

Walker, B. S. Effects of inserted questions on retroactive inhibition in meaningful verbal learning. *Journal of Educational Psychology,* 1974, *66,* 486-490.

Watts, G. H., & Anderson, R. C. Effects of three types of inserted questions on learning from prose. *Journal of Educational Psychology,* 1971, *62,* 387-394.

Winograd, T. Discourse. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

van Dijk, T. *Recalling and summarizing complex discourse.* Unpublished Mimeo, University of Amsterdam, March 1975.

van Dijk, T. Macro-structures and cognition. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

van Dijk, T., & Kintsch, W. Cognitive psychology and discourse. In W. U. Dressler (Ed.), *Trends in text linguistics.* Berlin—New York: DeGruyter, 1976.

CONNIE L. JUEL
University of Texas at Austin

ROBERT L. SOLSO
University of Idaho

EFFECTS OF ORTHOGRAPHIC AND PHONIC STRUCTURES ON WORD IDENTIFICATION

Word identification is a complex process which involves both the ability of the reader and the nature of the textual material. Much of the contemporary research in reading was stimulated by the seminal work of Cattell (1886) who demonstrated that subjects were able to report more accurately the contents of briefly presented visual displays if they were embedded in words rather than in unrelated strings of letters. Current investigation into the Word Apprehension Effect (WAE) has been instrumental in identifying several cogent elements of word and letter identification as well as stimulating research in the area of contextual effects of word/letter identification (Reicher, 1969; Krueger, 1975; Smith & Spoehr, 1974).

Of particular importance to the present study are a large number of perceptual experiments that, in general, show that letter/word identification is dependent on the degree to which the visual display conforms to the regularities of the English language. For example, words high in sequential redundancy are recalled better than those low in redundancy (Miller, Bruner & Postman, 1954; Rumelhart & Siple, 1974). Mason (1975) hypothesized that spatial redundancy (viz. positional frequency) can be used to "augment visual featural information in the identification of individual letters" (p. 147). In one experiment she had good and poor sixth grade readers search for the presence or absence of target letters and found that good readers were superior to poor readers only when the target letter was displayed in a highly redundant position.

In addition to orthographic factors influencing letter/word identification, other researchers have demonstrated the influence of phonic structure and letter-sound correspondences on word identification (e.g., Calfee, Venezky & Chapman, 1969; Juel, 1980). Thus, two factors have been identified as influencing letter/word identification (orthography and phonic structure). These two factors tend to be related, and in previous research it has been difficult (or impossible) to specify the influence of one independent of the other.

The present experiment was designed to further specify the effects of orthographic and phonic structure on word identification in reading. We were particularly interested in maintaining a *reading* task. Some reaction time (RT) studies, using a letter search task, have shown effects of orthographic structures (Mason, 1975; Massaro, 1979). However, it can be argued that a letter search task may represent processes not isomorphic to those involved in going from a printed word to its meaning. In the current study we employed both a lexical decision task and a task involving matching a printed word to a picture which illustrated its meaning. In both tasks we examined the effects of orthographic redundancy, decodability (phonic regularity), and word frequency on identification times. Word frequency was included as a factor to control for any differences that may occur in the processing of common or rare words. It was hypothesized that common words are frequently identified holistically and thus require less mediated processing than do the identification of uncommon words. Common words should, therefore, show less effect of redundancy and phonic patterns on identification times than would uncommon words.

We were also interested in processing differences that may occur between good and poor readers, as suggested by Mason (1975). Thus, grade level and reading ability were included as factors.

METHOD

The current study consisted of two conditions. In Condition I, subjects saw a slide of a word and two pictures — one picture was unrelated to the word and one represented its meaning. Subjects pressed a RT time key corresponding to the correct picture. The word varied on three dimensions: Frequency (high or low), decodability (easy or hard phonic patterns), and orthographic redundancy (high or low summed positional frequencies of individual letters). In Condition 2, subjects saw only a word (or a non-word) projected on a screen; the dependent measure was RT in deciding if the visual display of letters made a word or not.

Subjects

A total of 48 children enrolled in a summer reading tutorial program at the University of Texas at Austin participated in the study. (The program was promoted by the local school district as providing all levels of reading instruction — from remedial to enrichment.) Subjects were randomly drawn from all second through fifth grade children who scored between 85 and 130 on the *Peabody Picture Vocabulary Test* (IQ equivalents). (Mean score for Low Ability Readers=100; Mean score for High Ability Readers=107.) Subjects were drawn until 12 high ability upper grade level students (Grades 4 and 5), 12 low ability upper grade level students, 12 high ability primary grade students (Grades 2 and 3), and 12 low ability primary grade students were selected. Subjects were classified as high ability if they scored 6 months to 2 years above their grade level on both the *Wide Range Achievement Test* and an informal reading inventory. Subjects were classified as low ability if they scored 6 months to 2 years below their grade level on both the WRAT and IRI.

Materials

In Condition I, 48 four-letter target words were selected which represented all possible combinations of the experimental word factors (easy/hard decodable words; high/low frequency words; and high/low orthographically redundant words). Easy decodable words had vowels which corresponded to a commonly taught phonic generalization (CV̄CE, CV̄C, or V̄V) and had no consonant digraphs. Hard decodable words had vowels which did not correspond to the above generalizations and/or silent or consonant digraphs. Frequency was determined by the total word frequency count of Carroll, Davies, and Richman (1971). The mean frequency for high frequency words was 1645 per five million, with a range from 111 to 93,191. The mean frequency for low frequency words was 34 per five million, with a range from 2 to 88. Orthographic redundancy was based on summed positional frequencies of letters using the norms of Solso and King (1976). The mean summed frequency for redundant words was 71,125 per million, with a range from 60,753 to 100,089. The mean summed frequency for nonredundant words was 35,363 per million, with a range from 18,208 to 46,828. An example of a high frequency, hard decodable, nonredundant word is "soup"; and a high frequency, easy decodable, redundant word is "time."

Apparatus

Visual stimuli were displayed on a back projected screen by means of a carousel projector. Each display was shown by meaas of a remote control switch which simultaneously initiated a 1/100 second timer. A constant delay of 1 second occurred between the starting of the timer and the visual display. A reaction time device was made which consisted of two keys mounted on a wooden board each of which when depressed stopped the 1/100 second timer.

Procedure

Subjects were tested individually. Each subject sat in front of a table with one hand resting on the RT keyboard. The experimenter randomly chose one of the two conditions.

In the Word/Nonword condition the subjects were instructed to look at the string of letters on the screen and decide if it was a real word or a nonword. A real word was defined as one whose meaning they knew. A nonword was defined as a string of letters which did not mean anything. Before actual testing, subjects practiced depressing the RT keys until the process was automatic. In both the Picture and Word/Nonword tasks the sequence of slides were reordered periodically to offset any order effects.

If the Picture task was selected, the subject was instructed to read the word on the screen, choose the picture (1 of 2) which represented its meaning, and press the RT key which corresponded to that picture. Prior to the main experiment several practice slides were presented until response was automatic.

Design

Both conditions employed a mixed design in which variables were factorially manipulated. The between subjects factors were Grade Level (primary or upper) and Ability (High or Low). The within subjects word factors were Decodability (Easy or Hard), Frequency (High or Low), and Orthographic Redundancy (Redundant or Nonredundant). There were six replications of each word type (i.e., six easy decodable, high frequency, orthographically redundant words) in the Picture Task. Three of these replications for *each* word type were randomly selected from the Word/Nonword task. In addition, an equal number of nonwords were used. Half of the nonwords were orthographically legal (e.g., "dila"); half were orthographically illegal (e.g., rglz).

RESULTS AND DISCUSSION

The data were analyzed by repeated measures analysis of variance. Only correct responses were analyzed. The error rate was approximately 8 percent. A significance level of .01 was used. Significant main effects and means appear in Table 1. Interactions which reached significance are shown in Figures 1 through 4. Means in Table 1 represent RT's from the instant the stimuli appeared on the screen until the key press. It should be noted that a logarithmic transformation of the data was performed, but yielded no significant differences from analyses performed on the raw data.

TABLE I

ANOVA's and Means for Significant Main Effects

	Picture Task			Word/Nonword Task		
Source	F	(Error)	Means[a]	F	(Error)	Means[a]
Grade	15.2*	(.527)		16.9*	(.629)	
Grade 2, 3			2.80			2.24
Grade 4, 5			1.98			1.30
Ability	17.5*	(.527)		6.9	(.629)	
High			1.95	(not significant		1.47
Low			2.83	at .01 level)		2.07
Frequency	117.9*	(.172)		11.0*	(.193)	
High			2.08			1.52
Low			2.72			1.73
Decodability	22.2*	(.135)		48.0*	(.160)	
Easy			2.27			1.59
Hard			2.52			1.99

*$p < .01$, F(1,44)

[a] mean seconds

Picture Task

As shown in Table 1, both Grade and Ability had significant main effects, with older and/or high ability readers having faster RT's. Frequency was significant as expected, with faster RT's for high frequency words. As shown in Figure 1, there was a significant interaction between Frequency and Ability (F=11.6, $p < .01$). Low Ability readers appear to be exceedingly disadvantaged in identifying or accessing the meaning of low frequency words.

There was a significant main effect of decodability (Table 1). The meanings of words containing commonly taught phonic units are accessed faster than words lacking such elements. Apparently, young readers decode words into their constituent phonic units before accessing meaning. In other words, they were not using pictorial context plus initial consonants for identification.

There was no main effect of orthography. But as shown in Figure 2, there was a significant interaction between Orthography and Frequency (F=15.5, $p < .01$). As hypothesized, high frequency words often appear to be recognized holistically; whereas low frequency words require more mediated processing of internal orthographic characteristics. However, low frequency words, which are also low in redundancy, were identified *faster* than low frequency words, high in redundancy. We would have expected redundancy to have a facilitative effect. Instead, it appears to have a negative, confusability effect: Words which contain letters that are very common in their respective position within the word resemble other words and thus require more careful (and slower) decoding. A significant interaction between Orthography and Grade (F=7.7, p<.01; Figure 3) further indicated that older readers are most responsive to orthography and more "confused" by redundant words than young readers, who are relatively insensitive to redundancy. Apparently, considerable exposure to print is critical to developing a set for orthography.

Word/Nonword Task

As shown in Table 1, Grade had a significant main effect on RT to decide if a letter string was a real word; older readers were more adept at this task. Ability, however, was not significant at the .01 level (it was at the .05 level). Apparently, age and subsequent exposure to print are critical factors in determining "wordness".

Frequency had a significant main effect. High frequency words were more rapidly recognized as real words. However, the magnitude of this effect is considerably less than in the Picture condition. On the other hand, the magnitude of the significant main effect of decodability is considerably greater than in the picture task. There was also a significant interaction between Frequency and Decodability (F=12.9, p<.01; Figure 4). Decodability had little impact on high frequency words, again suggesting that such words are frequently processed holistically. However, easy decodable, low frequency words are reacted to considerably faster than hard decodable, low frequency words. Again, we have evidence that such words are being internally decoded. However, we find no significant main effect or interactions with orthography.

GENERAL DISCUSSION

Two distinct processes were disclosed in the present experiment. We find a different pattern of significance in the Picture task than in the Word/Nonword task, and overall, we find faster reaction times in the Word/Nonword task. We are sure in the Picture task that word meanings were accessed; we are doubtful word meanings were always accessed in the Word/Nonword task. Our instructions to the children in the Word/Nonword task were to decide that a word was a nonword by whether or not they knew its meaning. It appears, however, that children frequently made decisions *before* accessing meaning. That is, they made decisions about "wordness" based on presemantic characteristics, the most salient of which appeared to be the phonic structures in the word. We may have tapped a prerequisite stage to semantic identification; but this cannot be determined from the present study.

In general it appears that phonic structures are a critical factor in word identification for both good and poor elementary age readers. It also appears that certain orthographic characteristics play a supporting role in the identification of low frequency words. In particular, words with distinctive, nonredundant orthographic patterns are easier to identify than orthographically redundant words. However, only older readers were sensitive to this orthographic cue.

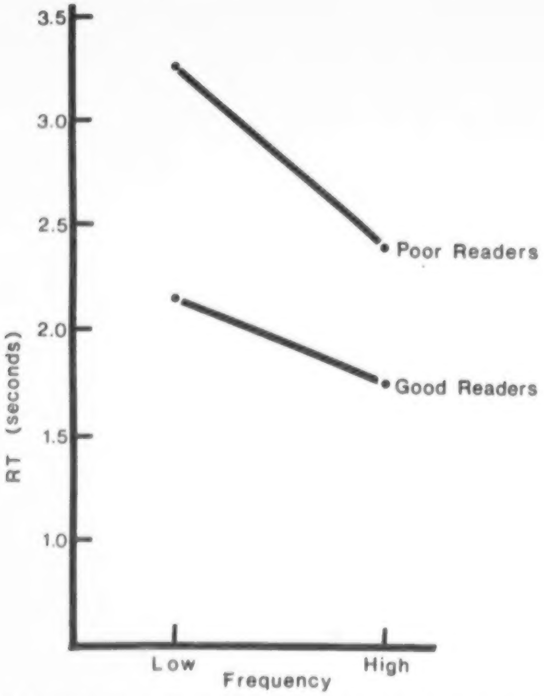


Figure 1: Interaction of Frequency and Ability in the Picture Task

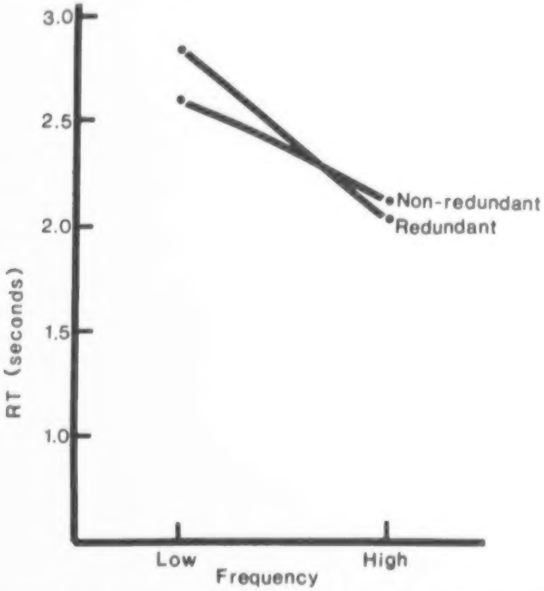


Figure 2: Interaction of Frequency and Orthography in the Picture Task

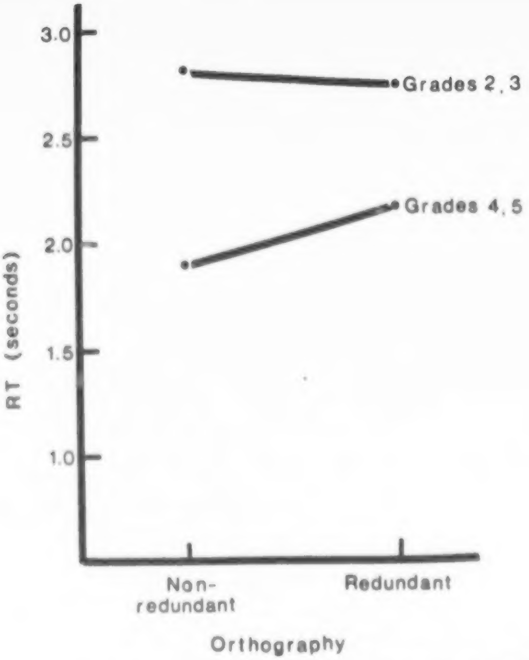


Figure 3: Interaction of Grade and Orthography in the Picture Task

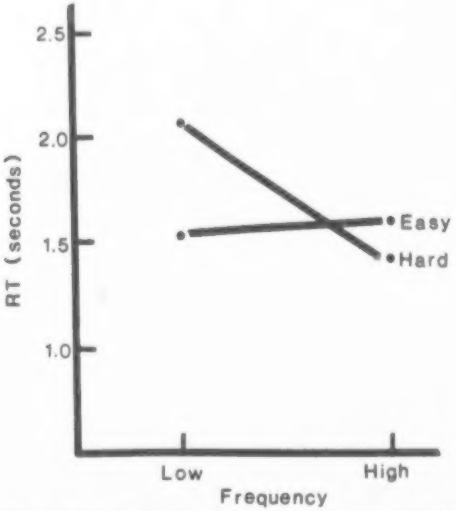


Figure 4: Interaction of Frequency and Decodability in the Word/Nonword Task

REFERENCES

Calfee, R. C., Venezky, R. L., & Chapman, R. S. *Pronunciation of synthetic words with predictable and unpredictable letter-sound correspondences* (Tech. Rep. 71). Madison: Wisconsin Research and Development Center for Cognitive Learning, 1969.

Carroll, J. B., Davies, P., & Richman, B. *Word Frequency Book*. New York: American Heritage, 1971.

Carroll, J. B., Davies, P., & Richman, B. *Word frequency book*. New York: American Heritage, 1971.

Cattell, J. McK. The time it takes to see and name objects. *Mind*, 1886, *11*, 63-65.

Juel, C. L. Comparison of word identification strategies with varying context, word type, and reader skill. *Reading Research Quarterly*, 1980, *15* (3), 358-376.

Mason, M. Reading ability and letter search time: Effects of orthographic structure defined by single-letter positional frequency. *Journal of Experimental Psychology: General*, 1975, *1*, 146-166.

Massaro, D. W. *Reading and orthographic structure*. Paper presented at the National Reading Conference, San Antonio, Texas, December 1979.

Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology*, 1954, *50*, 129-139.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimuli material. *Journal of Experimental Psychology*, 1969, *81*, 275-280.

Rumelhart, D. E., & Siple, P. Processes of reorganizing tachistoscopically presented words. *Psychological Review*, 1974, *81*, 99-118.

Smith, E. E., & Spoehr, K. T. The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Potomac, Maryland: Erlbaum Press, 1974.

Solso, R. L., & King, J. F. Frequency and versatility of letters in the English language. *Behavioral Research Methods and Instrumentation*, 1976, *8*, 283-286.

DOMINIC W. MASSARO
University of Wisconsin

READING AND ORTHOGRAPHIC STRUCTURE

Reading involves the integration of the visual patterns on a page of text with the reader's knowledge. Students of reading seek to understand the processes involved in evaluating the visual patterns, the nature of the knowledge sources, and how the knowledge sources are integrated with visual processing. It is generally accepted that sentential constraints in the form of semantic and syntactic information are utilized in reading. Although it is not as obvious, knowledge of orthographic structure (the constraints or redundancy of written language) is also utilized in letter and word recognition.

The utilization of orthographic redundancy is not necessary for accurate letter recognition, but it allows reading to be much more efficient. Even though the page of text appears to be clear to the reader, all readers have tunnel vision. Only 6 or 8 characters to the right and left of fixation are perfectly legible. In addition, the short duration of an eye fixation may not allow complete processing of all of the potentially legible letters. Orthographic structure supplements the relatively limited visual information available to the reader in a given eye fixation (Massaro, 1975).

Given that orthographic structure facilitates reading, it is important to describe the reader's form of knowledge of this structure. In previous work, Venezky and Massaro (1979) have distinguished between two categories of descriptions: statistical redundancy and rule-governed regularity. The first describes the reader's knowledge in terms of the frequencies of occurrence of letters and letter sequences in written text. According to the second description, the reader acquires certain rules about the nature of letter patterns in words. The rules can be described by phonological constraints in spoken English and scribal conventions in the transcription of spoken English into writing.

One method of testing various descriptions of the reader's knowledge of orthographic structure is to ask the reader to process novel letter strings. In our work (Massaro, Venezky, & Taylor, 1979), we begin with real words and make anagrams of them to create strings that vary systematically according to various descriptions. We have used two basic tasks in the study of the reader's processing of these letter strings. The first involves perceptual recognition of the letter strings whereas the second requires an overt judgment about the orthographic structure of the strings. In a protracted series of experiments, we have found large effects of orthorgraphic structure in both perceptual recognition and overt judgment tasks. Given these positive results, we were encouraged to assess how knowledge of orthographic structure related to reading ablity and how this knowledge develops over the time course of children learning to read.

Massaro and Taylor (1979) carried out a perceptual recognition task with college sophomores. A large introductory class in psychology was given the *Nelson-Denny Reading Test*. Students scoring in the top 15 percent were contrasted with their peers scoring in the bottom 15 percent of the sample. Although there was a large contribution of orthographic structure to perceptual recognition, this contribution was identical for good and poor readers. These adult readers appear to have mastered the utilization of orthographic structure in letter and word recognition even though they vary enormously in reading ability.

Three groups of sixth graders and one group of college students participated in a second experiment (Massaro & Taylor, 1979). Based on the comprehension subtest of the *Gates-MacGinite* reading test, the mean comprehension grade level for three groups of sixth graders was 5.0, 7.8, and 11.9, respectively. The effect of orthographic structure on perceptual recognition was 16 percent for adults, 10 percent for the good and average sixth-grade readers, and only six percent for the poor readers. Although this substantial interaction was not statistically significant, post hoc correlations between various descriptions of orthographic structure and performance on the 200 test items revealed significant differences due to reading level. A frequency-based description correlated .48, .46, .35, and .28 respectively for the adults, good, medium, and poor sixth-grade readers. A description based on rule-govered regularity also discriminated among the 4 groups of readers. These experiments provide some evidence that the utilization of orthographic structure may be related to reading ability but only for young readers.

In another study, Massaro and Hestand (1980) studied how the acquisition of knowledge of orthographic structure related to grade level and reading ability. Rather than using a perceptual recognition task, an overt judgment task was utilized. The subject is shown a pair of letter strings and is asked to choose the letter string that looks more like a word. This task provides information about the extent to which knowledge of orthographic structure is consciously available and capable of report. Previous experiments utilizing this task revealed that college readers discriminate among letter strings primarily in terms of rule-govered regularity rather than in terms of statistical redundancy (Massaro, Taylor, Venezky, Jastrzembski, & Lucas, 1980). Therefore, in each pair of items, one item did not violate any rules and the other had anywhere from one to four violations. First, second, and third graders at the end of the school year were tested. First graders averaged about 58 percent correct, second graders averaged about 69 percent correct and third graders averaged about 79 percent correct. Performance correlated .5 with grade level. We also asked to what extent performance on our orthographic structure test varied with reading level, as measured by the *California Achievement Tests*. Performance correlated .66 with reading level based on vocabulary and comprehension tests. Therefore, reading level accounted for 20% more of the total variance than that accounted for by grade level.

We cannot state from this research how important utilization of knowledge of orthographic structure is for good reading. However, the significant findings encourage us in the belief that it may be worthwhile to modify classroom practice to more directly facilitate the child's understanding of orthographic structure. There are some strong constraints in English orthography and these might be taught in a game format in the classroom (Massaro & Taylor, 1979). It is encouraging that the constraints in English orthography do not change significantly with the reading level of the text. We have found very high correlations among letter and letter-pattern occurrences in texts sampled from grades 3 to 9 (Carroll, Davies, & Richman,

1971) and texts for adult readers (Kucera & Francis, 1967). Some of the major constraints are differences in terms of where consonants and vowels occur in words. Vowel sounds are relatively infrequent in initial and final position in English words and therefore the reader can expect most words to begin and end with consonants. Overall constraints are greatest at beginnings and ends of words. Syllable boundaries in the middle of words permit almost any pair of letters to occur in medial positions. There are also major constraints on the possible consonant and vowel clusters and where they can occur in English words.

The games involve the child in determining which letters and letter patterns occur and where they occur in English words. In one game students are given a list of letter strings and asked to search for a particular letter or cluster of letters. If the child knows where letters and letter clusters normally occur, the search through the list will be much easier. An important aspect of the game is to discuss what rules of thumb are helpful in determining where letters and letter clusters might be found. In another game, the child arranges letters and letter clusters on a magnetic board to create new words. The goal is to create the words visually rather than orally. A game modeled after the overt judgment task is to categorize letter strings as possible or impossible words. Whenever a letter sring is judged to be impossible, the student indicates what is wrong with it and how it might be corrected. A popular game is a version of hangman; the student is told the number of letters of a mystery word and must guess what the word is. The participant guesses one letter at a time and is shown the position(s) in the mystery word of correct guesses. Playing these games, students might be made aware of many of the constraints in English orthography.

In summary, our work has encouraged our belief in the important contribution of orthographic structure in reading and learning how to read. Although phonics instruction is now generally accepted, little reading instruction directly teaches orthographic structure. Classroom programs incorporating instruction on orthographic structure might allow us to assess the role this knowledge plays in learning to read.

REFERENCES

Carroll, J. B., Davies, P., and Richman, B. *The American Heritage Word Frequency Book.* New York: American Heritage, 1971.

Kucera, H., & Francis, W. N. *Computational analysis of present day American English.* Providence: Brown University Press, 1967.

Massaro, D. W. (Ed.). *Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics.* New York: Academic Press, 1975.

Massaro, D. W. Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception & Performance,* 1979, *5,* 595-609.

Massaro, D. W., & Hestand, J. *Developmental relations between reading ability and knowledge of orthographic structure.* Unpublished manuscript, 1980.

Massaro, D. W., & Taylor, G. A. *Reading ability and the utilization of orthographic structure in reading* (Tech. Rep. No. 515). Madison, Wisc.: Wisconsin Research and Development Center for Individualized Schooling, University of Wisconsin, 1979.

Massaro, D. W., Taylor, G. A., Venezky, R. L., Jastrzembski, J. E., & Lucas, P. A. *Letter and word perception: The role of visual information and orthographic structure in reading.* Amsterdam: North-Holland, 1980.

Massaro, D. W., Venezky, R. L. & Taylor, G. A. Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General,* 1979, *108,* 107-124.

Venezky, R. L., & Massaro, D. W. The role of orthographic regularity in word recognition. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, N.J.: Erlbaum, 1979.

ACKNOWLEDGMENTS

The research reported in this paper was supported in part by funds from the National Institute of Education, Department of Health, Education, and Welfare. The contributions of J. Hestand, J. E. Jastrzembski, P. A. Lucas, G. A. Taylor, and R. L. Venezky are gratefully acknowledged.

JEROME A. NILES
Virginia Polytechnic Institute and State University

THE EFFECTS OF SELECTED TEACHER PROMPTING STRATEGIES ON ORAL READING PERFORMANCE

During normal reading, different types or levels of information such as visual feature, letter, word, syntactic and semantic are processed (Adams & Collins, 1979). The fluent reader artfully integrates the information from each level to extract the author's intended meaning. Rumelhart (1977) has suggested that this integration takes place simultaneously such that each level is contributing to the reader's overall purpose, which is the construction of meaning. How these levels are acquired is still a matter for argument and how the acquisition of integration is carried out is even more puzzling (Guthrie, 1973). Of interest to this study is how teachers, through instructional prompts, may affect beginning readers' oral reading behavior (miscues and comprehension) by attempting to focus the reader's attention on a specific level of information.

Learning to read is a process of discovering what is important on a printed page and what can safely be ignored (Smith, 1978). Children's success in making these discoveries has been well documented (Gibson & Levin, 1975). That instruction has a pronounced effect on how these discoveries are made has been demonstrated by Barr (1975). She found that when pronouncing words from a word list, students instructed in a skills model of reading with a whole word emphasis tended to substitute whole words, while students instructed in a phonics based program responded with more words that were orthographically similar. Thus, it seems reasonable to assume that specific teacher prompting strategies, which are part of an instructional program and direct the reader's attention to specific sources or levels of information in the text such as orthography or meaning, may also affect the reader's performance.

One of the most common strategies teachers have to direct a reader's attention is through feedback during oral reading. During oral reading sessions whenever the child makes an unexpected response to print, teachers may make a number of decisions. First, they must decide whether to call attention to the deviation or not. If yes, they must select the point at which they will draw the reader's attention to the deviation. Finally, they must choose the level to which they will direct the reader's attention such as orthographic, with a prompt such as "Sound it out," or semantic, "Does that make sense?".

Two previous investigations have shown that teacher prompts do have an effect on reading behavior. Pehrsson (1974) found that comprehension, based on free recall, suffered adversely when the teacher interjected prompts for the correction of word recognition errors. Similarly, Niles, Graham and Winstead (1977) found that recall of readers who read without correction was superior to that of readers who were prompted to correct their unexpected responses. Moreover, readers who received graphophonic prompts were more likely to make responses which were more graphically similar and changed the author's intended meaning. However, these two investigations compared only a graphophonic prompt from the teacher with an uncorrected condition which assumed the reader would make any necessary corrections. The present study examines the effects on reader behavior in two areas, miscues and comprehension, using the uncorrected prompt and the orthographic prompt as well as two additional prompting strategies. One of these strategies specifically directs the reader's attention to meaning by asking "Does that make sense?". The second prompt forces readers to make their own choice of level of attention as it simply asks the reader to repeat the sentence.

It is expected that the performance of participants under the orthographic prompting condition will differ both quantitatively and qualitatively from performance under the other prompting condition which either directly or indirectly focus the reader's attention on other levels of the text.

METHOD

Subjects

Forty beginning-of-the-year, third grade children were selected from a population of average readers whose percentile ranks on the reading subtest of the SRA Achievement Tests ranged between 38 and 81. The students were randomly assigned to one of four treatment groups, each containing ten members. Only 34 participants were eventually included in the analysis because of absences on more than one treatment day.

Procedure

On five consecutive days the participants read orally a different passage of approximately 300 words under their assigned treatment condition. The five passages, one for each day of the treatment and final testing, were of comparable difficulty and randomly assigned to each day. All passages were taken from the same third grade basal and verified to be of equal difficulty using the Fry readability formula (Fry, 1972). After reading the story under the assigned treatment condition, the participants were asked to tell the examiner all they could remember about what they had read. One additional prompt was given after the free recall was finished: "Can you remember anything else?". Following the free recall, a series of rote and transformed rote questions were asked.

The participants were randomly assigned to three examiners. No examiner worked with more than four participants from one condition. The first four days were considered training and the fifth day, with no change in procedure, was treated as the testing day. The entire session was tape recorded and only the data from the fifth day were used in the analysis.

Treatment

Uncorrected. Participants read without any substantive feedback from the examiners. If the readers paused for more than five seconds, the examiners encouraged them to make a guess and continue reading.

Graphophonic. When the participant made an unexpected response (e.g. horse for house) the examiner immediately (at the point of the response) called the reader's attention to the deviation by pointing to the word and prompting, "Look closely at this word," or when appropriate, "Sound this word out." Words such as *enough* received the first prompt and one such as *sleep* received the second prompt. The intent was to direct the readers' attention to orthographic information.

Semantic. When the participant made an unexpected response which changed the author's intended meaning (Goodman & Burke, 1972), the examiner prompted the reader by asking, "Does that make sense?". This prompt was typically interjected after the reader had completed reading the sentence or a complete thought within a complex sentence. The reader then decided what did not make sense and whether to regress or not.

Repeat. The participants in this condition were also prompted when they made an unexpected response which changed the author's intended meaning. After reading the complete thought in which the meaning was changed, the readers were prompted by the examiner to repeat that sentence or sentence part. The prompt was, "Would you repeat the sentence, please?". Unlike the semantic condition these readers were not asked if their response made sense.

Design and Analysis

A single factor design, prompting condition, was used which included four levels: uncorrected, graphophonic, semantic, and sentence repeat. The dependent measures were: 1) the participants' oral reading miscues, 2) free recalls, and 3) total number of questions answered. Miscues were examined for their graphic similarity, semantic acceptability, and meaning change (Goodman & Burke, 1972). Free recalls were scored for total idea units using a discourse analysis procedure which was developed by Taylor (1978) and based on a modification of procedures proposed by Meyer (1975) and Grimes (1975). Data were analyzed using an analysis of covariance procedure which held reading level constant. Tukey HSD procedures were used to test for significant mean differences where appropriate.

RESULTS

The readers' responses were evaluated using the *Reading Miscue Inventory* procedures (Goodman and Burke, 1972) which essentially uses a three point scale with yes and no representing the extremes. Table 1 displays the means and standard deviations for graphic similarity, semantic acceptability, and meaning change categories. This table presents the yes responses as well as the total no responses for each RMI category to give a more complete description of the readers' performance. In comparing performance on graphic similarity for the *no* dimension, for example, the sentence repeat readers apparently rely less on the graphic cue system than readers in the other three conditions. But when the *yes* dimension is examined, it can be seen that performance under the repeat condition looks more like performance under the uncorrected condition and the semantic condition.

While it appears that the graphophonic condition produced more responses which were graphically similar (*yes*) as well as graphically dissimilar (*no*), the analysis of covariance did not formally support these observations, $F(3,29)$ $p<.10$ and $F(3,29)$ $p<.21$ respectively. The covariance analysis for semantic acceptability verifies that there were significant main

TABLE I
Means and Standard Deviations for Percent of Graphic Similarity, Semantic Acceptability and Meaning Change of Miscues

		Yes			No	
		M	SD		M	SD
Graphic Similarity	U	49.3	12.3	U	16.9	10.1
	G	64.9	18.6	G	14.4	8.9
	S	50.5	27.2	S	19.3	11.9
	R	52.3	12.9	R	25.0	10.0
Semantic Accept	U	28	10.4	U	19.2	9.2
	G	16	12.4	G	37.7	15.2
	S	38.9	14.2	S	18.0	13.7
	R	31.1	15.0	R	12.9	5.9
Meaning Change	U	49.3	20.3	U	30.4	11.6
	G	67	10.5	G	20.5	12.8
	S	47.7	13.9	S	37.0	14.5
	R	44.1	15.9	R	32.1	13.7

U = uncorrected condition
G = graphophonic condition
S = semantic condition
R = repeat condition

effects for the two dimensions of the four prompting conditions, $F(3,29)=5.1$, $p<.01$ and $F(3,29)=9.07$, $p<.01$ for similar and dissimilar respectively. Post hoc analysis using Tukey HSD procedures indicated that the graphophonic condition rendered significantly fewer responses which were semantically acceptable than the semantic condition ($p<.01$). Moreover, the graphophonic condition generated more semantically dissimilar responses than all other conditions ($p<.01$). The meaning change category reflected the same pattern. That is, there was a main effect for both no change in meaning and significant change in meaning, $F(3,29)=3.21$, $p<.01$ and $F(3,29)=5.07$, $p<.01$ respectively. Further analysis using Tukey HSD procedures indicated that the semantic condition responses tended to change the author's intended meaning less than the graphophonic condition ($p<.01$). Additionally, the graphophonic condition responses changed the author's intended meaning more than the semantic condition and the repeat condition ($p<.01$).

For recall, no significant differences were found for the total idea units $F(3,29)$ $p>.20$. On the other hand, asking questions was marginally significant, $F(3,29)=2.73$, $p<.06$. The most disparate means representing correctly answered questions (maximum score of 8) for this measure were the uncorrected condition (3.67) and the repeat condition (5.25). The graphophonic condition (4.25) and the semantic condition (4.75) fell in between.

The amount of correction for each group varied. As expected the most often prompted group, the graphophonic condition, made the most corrections, 68% and the unprompted group made the least, 20%. The semantic condition successfully corrected 37% of their responses pointed out by the examiners and the repeat condition 54%.

DISCUSSION

In general, for the analysis of oral reading miscues, this study replicates the qualitative findings for miscues of a previous investigation (Niles et al., 1977) and supports the notion that teacher prompts can focus the reader's attention on different levels of information. Readers who were prompted with a graphophonic prompt produced less semantically acceptable responses which changed the author's intended meaning than readers who were not similarly prompted. The two additional prompting conditions of this study, semantic and sentence repeat, rendered results which were similar to the unprompted condition when compared to the graphophonic condition. No significant differences were found among the unprompted, semantic and sentence repeat conditions for semantic acceptability and meaning change. Previously found differences in free recall (Pehrsson, 1974, and Niles et al., 1977) were not replicated. In fact, no significant differences were observed across any of the four conditions.

While the graphic similarity of reader miscues across all the conditions did not differ significantly, there is evidence of a trend that graphophonically prompted readers produced more responses which were graphically similar (64%) as compared to an average of (51%) for the other three conditions. This trend may not have achieved statistical significance because of the importance of the orthographic level of information to beginning third grade readers. This importance may be intrinsic to the learning to read process or it may be an instructional artifact (Barr, 1975). In either case, the performance of the participants in using graphic information was uniformally higher than performance for the semantic level information. Thus, the graphophonic condition represents an escalated performance on an already heavy use of this particular level of information, yet not high enough to be significantly different. Interestingly, all three other conditions were within three percentage points of each other.

All beginning readers in this study demonstrated that they were using different levels of information to formulate their reading responses. However, those that were prompted by being asked if their reading made sense or by being asked to repeat sentences in which meaning was changed relied more heavily on meaning as part of their reading strategies. The graphophonic readers on the other hand, apparently heeded the teacher's prompt to pay closer attention to the orthographic structure of the word as their responses reflected a more extensive use of the graphic input than the meaning input.

The predicted quantitative difference in recall was not observed for the retellings. Very few idea units were recalled by any of the readers. The highest score was 37 out of 143 idea units and the average was approximately 13.5. This unexpectedly low recall performance may have been due to the expository nature of the materials and/or the elicitation procedure for this study which used only one additional query after the recall was completed. The fact that questions did reflect a marginal effect for prompting condition deserves comment. The sentence repeat condition yielded the highest performance and the uncorrected condition the lowest. It is plausible that the added advantage of rereading parts of the passage gave the repeat readers an advantage in recall. However, it is also possible that the request to repeat may have acted as an attentional cue for the reader that "something was wrong with the meaning I just made." In fact, by the fourth treatment several readers would beat the examiner to the prompt by saying when they made a perceived meaning change, "Oh, do you want me to read that sentence again?" The diminished performance by the unprompted readers may have been due to their tendency to get occasionally untracked in their search for meaning. The unprompted readers 20% average correction rate for responses which changed meaning suggests that they kept moving rather than regressing to build a more appropriate meaning. The higher correction rates of the other conditions may have helped in improving questioned recall; yet the highest rate, which was obtained by the graphophonic condition, rendered the second lowest performance.

The instructional implications from this study come from the conclusion that, while beginning readers do use various levels of informational input, the teacher, through prompts during oral reading, can focus the reader's attention more extensively on one level or another. Since there is evidence that teachers apparently do differentiate their prompting strategies based on reading ability (Allington, in press), it is important that they be aware of what type of prompts are used and what concomitant effects the prompts are having on the reader's performance. Assuming that normal reading is an interactive process, teachers of beginning readers can use various types of prompts as well as provide ample opportunities for self-prompting to help readers attend to and successfully integrate all levels of information needed to reconstruct the author's intended meaning.

REFERENCES

Adams, M., & Collins, A. A schema-theoretic view of reading. In R. Freedle (Ed.), *New directions in discourse processing* (Vol. 2). Norwood, New Jersey: Ablex Publishing Co., 1979.

Allington, R. Teacher interruptions during primary grade oral reading. *Journal of Educational Psychology*, in press.

Barr, R. The effect of instruction on pupil reading strategies. *Reading Research Quarterly*, 1975, *10*, 555-582.

Fry, E. *Reading instruction for classroom and clinic.* New York: McGraw Hill Book Company, 1972.

Gibson, E., & Levin, H. *The psychology of reading.* Cambridge, Mass.: The MIT Press, 1975.

Guthrie, J. Models of reading and reading disability. *Journal of Educational Psychology*, 1973, *65*, 9-18.

Goodman, Y., & Burke, C. *Reading Miscue Inventory.* New York: The MacMillan Company, 1972.

Grimes, J. *The thread of discourse.* The Hague, Netherlands: Mouton and Company, 1975.

Meyer, B. *Organization of prose and its effects on recall.* Amsterdam: North Holland Publications Co., 1975.

Niles, J., Graham, T., & Winstead, J. *The effects of teacher corrections on children's oral reading performance.* Paper presented at the 22nd Convention of the International Reading Association, Miami, 1977.

Pehrsson, D. How much of a helper is Mr. Gelpher? *Journal of Reading*, 1974, *19* (8), 617-621.

Rumelhart, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6). London: Academic Press, 1977.

Smith, F. *Understanding reading.* New York: Holt Rinehart & Winston, 1978.

Taylor, B. *A comparison of good and poor readers ability to read for meaning.* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, 1978.

CHARLES TEMPLE
University of Houston Victoria Campus

LEARNING TO SPELL IN SPANISH

YUTS A LADE YET FEHEG AD HE KOT FLEPR ("Once a lady went fishing and she caught flipper." From C. Chomsky, 1975)
EL PERRO CUADO ETABA HICITO CE ALIMETABA DE LA MADRE ("El perro cuando estaba chiquito se alimentaba de la madre." Collected by the author)

In recent years a few researchers have made note of the phenomenon of "invented spelling" in young children. They have demonstrated that many children of preschool age are capable of producing written messages using words that are generated through an original system of orthography—using clear patterns for combining letters to make words; words that obviously have not been imitated from adults. Read was the first to bring a linguistic analysis to bear on invented spelling. Henderson and his students—Beers, Zutell, Templeton, Gentry and Stever—have carried the work a lot further, especially with school-aged children.

By now we can assert that invented spelling is a widespread phenomenon, and we can list some of its central features:

a) It is systematic: at any one time a child appears to have a pattern to the way he (she) spells words. A young speller employs a small subset of all possible letters that could represent parts of words, and he makes the same choices repeatedly and consistently.
b) It is based initially upon phonetic principles—upon associations between sounds in words and graphemes. Later the spelling of most children moves away from a direct phonetic representation to honor other features of standard spelling, such as conventions for marking vowels, grammatical features, morpheme boundaries, and eventually a whole (as yet little explored) set of lexical relationships among words. But in the early stages children's invented spelling is phonetic, and it involves a surprisingly sensitive set of phonetic judgments.
c) Even though early spelling is so clearly divergent from standard spelling as to appear bizarre to adult eyes, nevertheless, the major principles of invented spelling are virtually identical among children at the same developmental stage. A child living in Boston and a child living in South Texas will often invent the same spellings for the same words even though these spellings look not at all like standard form.

It is this last aspect of invented spelling—its widespread occurrence—that has especially intrigued the present writer. Does the occurrence of children's invented spelling strategies generalize to languages other than English? The writer undertook to see if he could elicit invented spelling from children who are native speakers of Spanish. Any non-standard spellings that the children produced would be studied to see if they seemed to have been generated by principles similar to those that had been observed in the invented spellings of English-speaking children.

Spanish was chosen as the language to be investigated in this study for two reasons. First, its writing system is characterized by a markedly greater degree of phoneme-to-grapheme regularity than English. in fact, among alphabetic languages, Spanish and English are often said to occupy different ends of the spectrum of abstract/direct phoneme-to-grapheme representation (cf. Read, 1975; Laubach & Laubach, 1960; Gelb, 1956). A second reason for choosing Spanish was that the investigator had often heard Spanish speakers claim they had no trouble learning to spell in their language, and that they had difficulty remembering their teachers' efforts to teach them to spell. This is certainly different from the typical experience of English-speakers with their own writing system.

PROCEDURE

The subjects chosen were fifteen first graders, thirty second graders, and thirty third graders attending a parochial school in the Dominican Republic. The writer composed a list of twenty words to have the children spell. The words were selected according to three criteria: first, they were in the spoken vocabulary of the subjects (as attested to by their teachers); secondly, they were not on any of the children's spelling lists; and thirdly, they embodied a cross section of the types of phoneme-grapheme correspondences found in the Spanish language.

A teacher in the school composed sentences to illustrate each of the words on the spelling list. When the list was tested, the subjects' respective classroom teachers read the words and the sentences, and encouraged the children to spell the words as they thought they might be spelled, without fear of being penalized if they made errors. The writer tape recorded the testing sessions, and later had several children from the classes read the test words, the sentences, and the names of the letters of the alphabet into the tape recorder. These were used later to aid in the interpretation of the results.

While preparing the spelling list for testing, the researcher made a list of all of the phoneme-grapheme units that were tested in the study. After testing, the students' papers were examined and their responses for each phoneme-grapheme unit were tabulated. Subsequently, a comparison was made among the responses of first, second and third graders for each phoneme-grapheme unit.

An aside about Spanish orthography

For the purposes of analysis, the phoneme-grapheme units in the Spanish word lists were divided into five categories; depending on whether they were: *intuitable, conventional, ambiguous, phonologically variant,* or *morphemic.*

Intuitable units are those in which the sound of the grapheme is closely related to the phoneme it represents. All monophthong or pure vowels are intuitable in Spanish, and so are most consonants when they occur at the beginning of a word.

Conventional units are spelling units in which the phonemes and graphemes are joined by a convention or rule, which is not intuitable. CH, and LL are conventional spellings for the phonemes /c/ and /j/, respectively, as in *choque* and *calle.*

Ambiguous units are phonemes which can be represented by more than one grapheme. For example /s/ may be spelled by C, S, or Z. C has this sound only before front vowels; but S or Z (rarer) may have this sound before any vowel. B and V offer a more interesting example. Linguists find there is no stable difference between the sounds represented by either letter, and they point out that there are no minimal pairs of words distinguished by a difference between /b/ and /v/. Teachers and the Real Academia Espanola insist that there *should* be a difference in pronunciation, and an overprecise /v/ is taught for V. This writer observed, however, that in casual speech, the teachers themselves did not observe the distinction.

Phonologically variant units are phonemes which are changed (usually deleted) when they occur in certain environments in a word. Continuant consonants (/s/, /m/, /r/, /l/, and /n/) which occur before stops are often deemphasized in speech, since most of their features become redundant with those of the following consonant. Examples in Spanish are *triste, verdad.* Word final consonants often are deleted in rapid speech, giving [utéh] for *usted,* and [verdá] for *verdad.*

Morphemic units are phonemes and graphemes whose presence serves as an indication of the syntactic function of the word to which they are attached. In *tienen*, for example, the final N indicates that the verb is third person plural. In *libertad*, the D is part of the nominalizer *-ad*, corresponding to -ty in English. Morphemic units are often word final consonants. Because they are subject to the phonological influences and are often deleted in casual speech, comparing their rate of inclusion with that of non-morphemic word-final consonants constitutes a test of whether a child is sensitive to morphemic considerations in spelling.

RESULTS AND DISCUSSION

The results of the analysis of the children's errors is tabulated below. The children's responses on each spelling unit will be discussed separately.

These vowels occurred in words other parts of which were often misspelled. The vowels themselves were spelled correctly a high percentage of the time, and there appeared to be no particular pattern to the errors that these vowels did invoke.

Conventional consonants showed a far lower correct rate than single-letter, non-ambiguous consonants. There was also a definite pattern to the errors they evoked. For the /c/ phoneme, conventionally spelled CH (called /se ace/) the letter H appearing alone was the most frequent error, accounting for 17% of the errors at first grade and 9% at second. This error can be explained by the letter-name spelling strategy. If one were searching for a single letter of the alphabet whose name contained most closely resembled the phoneme one wished to spell, then the letter H (called /ace/ in Spanish) would be the best bet, since it does indeed contain the alveolar

TABLE I

Spelling of Pure Vowels (Intuitable Units)

Test Grapheme		Grade 1			Grade 2			Grade 3		
		Response Grapheme			Response Grapheme			Response Grapheme		
E [e]	Spellings:	E	A	0	E		A	E		U
"pecho"	Number:	48	5	3	264		2	258		1
	Percent:	86%	9%	5%	99%		1%	99.9%		.01%
O [o]	Spellings:	0	A		0		A	0	A	U
"pecho"	Number:	43	1		366		1	349	2	3
	Percent:	98%	2%		99.8%		.02%	98%	.05%	.08%
A [a]	Spellings:	A	E		A			A		0
"cara"	Number:	33	2		442			351		1
	Percent:	94%	5%		100%			99.8%		.02%
I [i]	Spellings:	I			I	IE	U	I		
"triste"	Number:	24			175	1	1	165		
	Percent:	100%			99%	.05%	.05%	100%		
U [u]	Spellings:				U	A	0	U	A	
"ayudan"	Number:	No Exemplar			35	1	1	32	1	
	Percent:				95%	2.5%	2.5%	97%	3%	

TABLE II

Conventional Consonants

Test Grapheme		Grade 1			Grade 2			Grade 3	
		Response Grapheme			Response Grapheme			Response Grapheme	
CH [č]	Spellings:	CH	H	HC	CH	H	LL	CH	LL
"pecho"	Number:	21	5	4	118	12	5	128	2
"choque"	Percent:	70%	17%	13%	87%	9%	4%	98%	2%

affricate sound of /c/. Even the reversed spelling HC might be attributed to this strategy: if one sensed that the sound element one were looking for were contained in the letter H, and if one had been taught that the phoneme were conventionally represented by the letter C and H together, it would be defensible to put H—the really important element—right up in first position. Examples were HOCE or HCOCE for *choque*.

At second and third grade there was interference from another conventional spelling, LL. In Dominican Spanish, the /j/ phoneme is very strident pronounced, so that a slight affrication is sometimes sensed. /č/ and /j/ sound sufficiently alike to become confused in spelling. An example is LLECE for *cheque*.

Ambiguous consonants caused the most difficulty. In general the children appeared to choose one letter to represent a phoneme and stick to it regardless of the word. Note that with the spelling of V for /b/, the children's error rate actually increased with grade level so that by third grade two out

TABLE III

Ambiguous Consonants

Test Grapheme		Grade 1		Grade 2				Grade 3					
		Response Grapheme		Response Grapheme				Response Grapheme					
B [b], [ƀ]	Spellings:	B	V	B	V	D		B	V	P	D	F	G
"bano"	Number:	9	14	109	23	1		115	4	1	1	1	1
"bombon"	Percent:	39%	61%	82%	17%	1%		93%	3%	1%	1%	1%	1%
V [b], [ƀ]	Spellings:	V	B	V	B	D	P	V	B	D			
"verdad"	Number:			40	25	1	1	21	36	1			
	Percent:			60%	37%	1%	1%	36%	62%	1%			
C [s]	Spellings:	No Exemplar		C	S			C	S				
"cerrar"	Number:			10	21			18	14				
	Percent:			32%	68%			56%	44%				
S [s]	Spellings:	S	C	S				S					
"silla"	Number:	16	1	35				32					
	Percent:	94%	6%	100%				100%					

of three spellings for this consonant were incorrect. A clear preference for the letter S as a spelling for /s/ is also evident. Examples are VANO (*bano*), SERRAL (cerrar).

TABLE IV

Pre-Obstruent Continuants (Phonologically Variant Units)

Test Grapheme		Grade 1			Grade 2				Grade 3			
		Response Grapheme			Response Grapheme				Response Grapheme			
S/T	Spellings:	No Exemplars			S	Ø	R	C	S	Ø	P	
"triste"	Number:				39	26	1	1	56	8	1	
	Percent:				58%	29%	1%	1%	86%	12%	1%	
R/D or T	Spellings:	R	Ø	L	R	Ø	L		R	Ø	L	NN
"verdad"	Number:	10	7	3	62	22	15		82	6	6	1
"libertad"	Percent:	50%	35%	15%	62%	22%	15%		86%	9%	9%	1%
M – /B	Spelling:	No Exemplars			M	Ø	N	L	M	Ø	N	
"bombon"	Number:				5	5	20	1	12	3	17	
	Percent:				16%	16%	65%	3%	38%	9%	53%	
N – /C	Spelling:				N	Ø	M	L	N	Ø	M	L
"hinchado"	Number:				44	3	4	9	58	2	2	1
	Percent:				73%	5%	7%	15%	92%	6%	6%	1%

Whereas English has many cases where a continuant consonant occurs before an obstruent within a single syllable, Spanish does not. At the junction of two syllables, however, this combination may occur. When it does, the pronunciation of the continuant is indistinct. Often the subjects omitted the continuant in such cases, but sometimes they substituted other continuants. Examples are: TRITE (*triste*), ILHADO (*hinchado*).

The subjects omitted final consonants at a high rate or alternatively they often substituted another consonant. Whether or not a consonant was part

TABLE V

Word-Final Consonants (Morphemes and Others)

Test Grapheme	Second Grade			Third Grade		
	Correct	Omit	Substitute	Correct	Omit	Substitute
D	5	61	3	32	30	2
"verdad"	7%	88%	13%	50%	37%	3%
N_1	19	11	0	29	2	1
"bombon"	63%	37%		91%	6%	3%
N_2	41	20	3	46	12	1
(verb marker)	64%	33%	5%	78%	20%	2%
"tienen"						
R	15	7	11	12	1	20
(verb marker)	45%	21%	33%	30%	3%	61%
"cerrar"						

of a morpheme does not appear to have made appreciable difference. Phonetic features of certain consonants made them more salient than the morphemic status of others. Examples: LIBERTA (*libertad*), SERRAL (*cerrar*).

CONCLUSIONS

The children in this study, all native Spanish speakers, apparently had not been taught to spell the target words. All of them plunged in and attempted to spell words anyway. They clearly used generalizations or concepts about spelling as a system to generate the spellings. Even when they guessed, they rarely guessed wildly; the children as a whole for any given word produced a very small subset of all the possible variations of letters that might have been produced to spell the word.

The subjects could and did take advantage of many close correspondences between phonemes and graphemes in the Spanish writing system to spell many words correctly. However, there were many instances where these correspondences were not close, either because of features of the writing system itself that deviated from the "phonetically regular" norm or because pronunciation sometimes was relaxed. In both such cases, Spanish-speaking children resorted to the same inventive spelling strategies English-speaking children regularly employ in such cases:

—they sought a match between the name of the letter of the alphabet which was closest in place and manner of articulation to the phonemes they wished to represent;

—they observed a kind of threshold criterion for the need to represent a phoneme; if it was very clear, they put it in. If not, they left it out, lest they distort the word by overspecifying its sounds.

This study indirectly demonstrated the shortcomings of an orthography based largely on phoneme-grapheme correspondences. In Spanish, a language highly touted as having one of the world's most regular writing systems, sounds alone are not an adequate guide to correct spelling. The subjects in the study demonstrated a clear need to attend to other sorts of information: morphemes, for example.

Whether older children would demonstrate an awareness of non-phonetic information is not known. Teachers in universities where Spanish is spoken sometimes complain of numerous spelling errors that have a phonetic base (Marcos Marin, 1975). Interestingly, at least one authority promotes the study of etymology as a corrective influence on spelling (Alonso, 1965). That good soul is clearly at odds with the Spanish dictionary writers, who for years have been actively removing etymological vestiges from the spellings of words, all in the name of "el fonetismo", the regularity of phoneme-grapheme correspondence.

ACKNOWLEDGMENT

The author is deeply indebted to Sra. Ana Maria de Cid, Sra. Melania Harris de Lynch and the first, second and third graders at Colegio Todos Los Santos, La Romana, D. R., for their generous and enthusiastic cooperation with this study; and also to Rev. Wm. Skilton for related reasons.

REFERENCES

Alonso, M. *Apuntes de ortografia.* Merida, Venezuela; Universidad de los Andes, 1965.

Beers, J., & Henderson, E. A study of developing orthographic concepts among first graders. *Research in the Teaching of English,* 1977, *11,* 133-148.

Chomsky, C. Invented spelling in the open classroom. *Word,* 1971, *27,* 499-518.

Dalbor, J. *Spanish pronunciation: Theory and practice.* New York: Holt, Rinehart and Winston, 1969.

Gelb, I. *A study of writing.* Chicago: University of Chicago Press, 1956.

Gentry, R. *A study of orthographic strategies of beginning readers.* Unpublished doctoral dissertation, University of Virginia, 1977.

Laubach, F., & Laubach, R. *Toward world literacy: The each-one-teach-one way.* Syracuse, New York: Syracuse University Press, 1960.

Marcos Marin, F. *Linguistica y lingua espanola.* Madrid: Editorial Cincel, 1975.

Read, C. *Children's categorization of speech sounds in English.* Urbana, Illinois: National Council of Teachers of English, 1975.

Temple, C. *An analysis of spelling errors in Spanish.* Unpublished doctoral dissertation, University of Virginia, 1978.

Templeton, W. *An awareness of certain aspects of derivational morphology in phonology and orthography among sixth, eight and tenth graders.* Unpublished doctoral dissertation, University of Virginia, 1976.

Zutell, J. *Spelling strategies of primary school children in relation to the Piagetian concept of decentration.* Unpublished doctoral dissertation, University of Virginia, 1975.

ANNE D. FORESTER
Camosun College

USING A PROCESS MODEL AS A GUIDE TO COLLEGE/ADULT READING INSTRUCTION

INTRODUCTION

A review of the literature of college/adult reading reveals a wealth of well-designed instructional procedures and programs. Many of them are the result of careful attention to the needs of changing student populations. Yet few move beyond *ad hoc* problem solving and a focus on the development of discrete skills. There are programs based on the tenets of behavior modification (Himmelstein & Greenberg, 1974; Hornstein, 1970; Rankin, 1973; Reid, 1971) and on principles of humanistic psychology (Greene, 1974; Paul, 1974; Sherk & Manzo, 1973; Wilson, 1972). But there appear to be few programs designed specifically to consider the processes used by the effective adult reader.

There is little disagreement about the dynamic nature of reading and the fact that development continues beyond the acquisition of a set of skills. The effective reader continues to expand his ability to deal with texts of varying style and content. But in order to accommodate novel materials he makes recurring regressions to lower levels of functioning. At this point there does not appear to be a model that describes the processes the effective reader uses when dealing with unfamiliar or difficult texts. The purpose of this paper is to propose such a model and to discuss its implications for instruction.

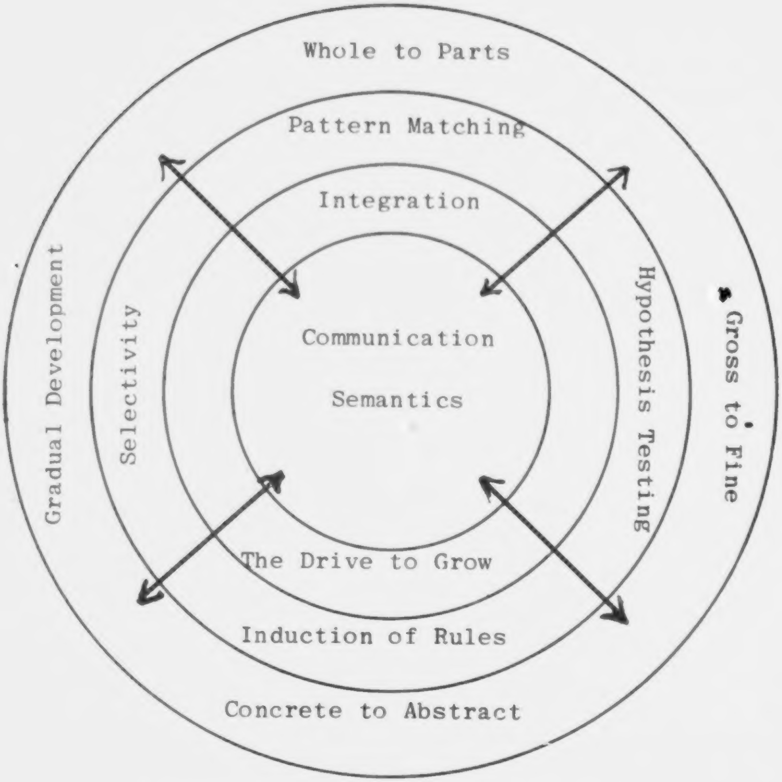
A PROCESS MODEL OF ADULT READING

Each time a reader encounters texts in styles or on topics unfamiliar to him he must internalize new linguistic and cognitive structures. In other words, he is faced with a task similar to learning a language. Both the child learning his mother tongue and the adult plunged into a foreign milieu must formulate new concepts and internalize the linguistic structures to describe them. The reader attempting to learn the mini-language of an unfamiliar text is faced with the same task. The assumption here is that he will use processes similar to those of the language learner to gain fluency. Accordingly, in the absence of case studies of adult reading development, the literature of psycholinguistics and of cognitive psychology has been used to derive the process model of adult reading presented here. (See Figure 1).

Communication stands at the heart of the model and interacts with all of its facets. However, before fluency is achieved, the learner moves through a number of levels. Depending upon his prior experience with the topics and language of the text he will linger at the periphery or move quickly to the induction of rules and the integration of the various aspects.

The lowest level constitutes a data gathering stage which is marked by initial gross performance. Reading is not a process of piecing together an

FIGURE 1


Whole to Parts
Pattern Matching
Integration
Gradual Development
Selectivity
Communication
Semantics
Hypothesis Testing
Gross to Fine
The Drive to Grow
Induction of Rules
Concrete to Abstract

A Process Model of College/Adult Reading

intricate mosaic of many individual pieces. Instead, it is akin to the development of a photographic plate where gross outlines emerge first and fine details develop later. The general topic and organization are perceived before there can be an appreciation of the subtleties of the text. If the reader has little or no experience with the material, he will search for concrete data in his own experience or the text before moving to more abstract levels of processing. At times, he may have to consult texts written in more concrete terms before being able to deal with abstract concepts of a field.

Instructors in reading centres are well aware of the data gathering stage of reading. They have observed how meaning accrues gradually as their students struggle with unfamiliar reading tasks. In literary analysis, the inexperienced reader moves from a retelling of the plot to descriptions of characters and events before he is able to engage in inferential reading or the analysis of theme or style. When dealing with abstract texts, the reader attempts to relate the information to his prior, more concrete knowledge and translates the unfamiliar terminology into more familiar language.

As in language learning, the move to higher levels of processing and the induction of rules requires active practice on the part of the learner.

Patterns of language and thought are internalized gradually. As they are consolidated, the reader becomes more effective in projecting ahead and selecting the right cues for effective reading. The presentation of rules or theorems cannot replace the processes of assimilation and accommodation in establishing the cognitive and linguistic schemata that will make for fluency.

The full integration of the linguistic and conceptual structures of a text constitutes the final move toward fluency. As the reader joins the components of the text and integrates them into his own conceptual framework the need to translate disappears and direct communication becomes established. This final step toward fluency is generally based on the learner's own initiative. Instruction may aid the process, but only the learner can fit the new structures into his own cognitive network.

While the independent, effective learner uses the holistic processes described here, reading instruction frequently runs counter to the move from the whole to parts and from the concrete to the abstract. Attention to the fine details of the text is stressed and accuracy rather than gross performance is expected on reading tests requiring a one-time perusal. Abstract theories and rules are presented to guide students in dealing with components of reading. If instruction is to foster the effective functioning of the learner's own processes, a shift may be needed in the overall approach to program development.

HOW TO FOSTER READING DEVELOPMENT

The recurring regressions the reader makes to accommodate new materials have been likened here to learning mini-languages; hence, the conditions that foster language learning in non-institutional settings are of interest. The child acquiring his mother tongue is eminently successful, and second-language learners immersed in a foreign milieu generally attain better fluency than their school-trained counterparts. In both instances, the data gathering stage is well supported by the learning environment. The learner is surrounded by whole language in concrete settings. No one expects initial accuracy of performance, and the learner is left to evolve his skills gradually.

In these non-institutional settings the teaching-learning process rests on three prime components: 1) The provision of models to emulate; 2) extensive practice on the part of the learner; 3) feedback that acknowledges the content of the learner's message and expands its form. The climate for learning is one of openness and trust. The learner feels safe to test his evolving skills, and his teachers display confidence in his ability to upgrade his initial efforts. Ideally, a learning centre should provide this type of climate and base instruction on the three prime components cited above.

Providing Models to Emulate

Reading instruction in institutional settings is generally based on explaining discrete skills. Without having the overall picture clearly in mind, the learner attempts to piece together a highly intricate mosaic. While individual skills may be acquired, learners often fail to move to the level of pattern formation, the induction of rules, or the integration of the materials read. In order to establish a solid data bank to serve as a foundation for higher levels of functioning, the learner needs models of behaviours that present an overview of what is aimed for.

Using the students' own textbooks or closely related materials at a lower, more concrete level, instructors can demonstrate reading and study behaviours incorporating the allied skills of writing, listening, and reasoning. For some students, fluent oral reading of poetry and prose may be essential as a guide to effective reading. Audio tapes can supplement personal demonstrations. In conjunction with textbook study, the art of formulating questions requires many concrete examples. It simply is not enough to tell students about turning headings or topic sentences into questions. Thinking aloud while moving through a chapter with students will serve as a model for in-depth analysis and the sifting of ideas. Similarly, literary analysis requires that the instructor reveal how or why he reaches certain interpretations.

Writing at all levels lends itself well to modelling. Here, as with reading, it becomes important to demonstrate that gross processing precedes fine discrimination in order to disspell the notion that a once-over-lightly treatment is all that is required. Students need to become aware that instructors, too, refine their work gradually in both reading and writing. To aid the modelling process, samples of essays and reports including rough notes and early drafts should be on file in the learning centre. Examples of varying types of sentences, paragraphs, and organization will serve students far better than the perennial exercises in classifying and/or correcting faulty writing.

Content-area instructors, too, can serve as models in the learning centre. There is no suggestion here that reading specialists be knowledgeable in all areas. In fact, all they need to do is model learning behaviour. If a text is difficult, the modelling will reveal the groping for familiar or concrete clues or patterns to aid the gradual move toward full comprehension. Peer models can supplement the work of the specialists.

Making Room for Practice

Instructors are well familiar with students' complaints about having to read chapters several times. The need for sustained practice and gradual improvement so thoroughly ingrained in athletic activities is not perceived when it comes to reading, writing, or thinking skills. To be effective, a learning centre must provide a climate that makes it safe for students to practice their emerging skills over a period of time. Instructors need to reassure students that repetition is not a matter of slow or faulty progress but a vital part of all learning. The need for active involvement of the learner in the formulation of cognitive structures is well established for early learning. Somehow it seems to be lost on college students and needs to be inculcated.

Providing Expanded Feedback

In language learning the focus rests squarely on communication. Parents or peers will acknowledge the content of the initial haulting efforts of the learner and at the same time provide feedback that expands the incomplete or incorrect statement. Thus the teacher reinforces communication and provides a model of correct forms of speech. In educational settings the opposite course is frequently followed. Feedback usually refers to faulty structures, spelling, or punctuation and often fails to acknowledge the content of the student's work.

When using the process model as a guide to instruction, there is a shift to content-centered responses in discussions of student's interpretations of texts. The use of expanded feedback lends itself well to both textbook reading and literary analysis. As the student provides brief or incomplete statements about the text, the instructor acknowledges the message and expands its form. In such an exchange the instructor, like the parent in the home, stays just enough ahead of the learner to aid the upgrading process without making too large a stride for the learner.

SUMMARY

The process model of reading suggests that even the effective reader makes recurring regressions to lower levels of functioning when faced with unfamiliar topics or texts. The gradual formulation of new linguistic and cognitive structures necessitated by the task of reading is likened to the acquisition of mini-languages and the model suggests that the learner moves through a number of stages before gaining fluency. To foster the smooth progression from initial gross processing to full comprehension and fluency, the reader needs conditions that allow the full functioning of his own cognitive and linguistic structuring. The suggestions provided here for creating a learning climate require a shift in philosophy more than in method or selection of materials. The tripartite approach of modelling, practice and expanded feedback can be fitted into most existing programs. The important aspect is to have — and to create for the student — an awareness of the gradual nature of learning generally and reading comprehension in particular.

REFERENCES

Greene, D. M. Toward personal growth through reading. In M. P. Douglass (Ed.), *Reading, thought and language.* Thirty-eighth Yearbook of the Claremont Reading Conference. Claremont, Cal.: The Claremont Reading Conference, 1974.

Himmelstein, H. C., & Greenberg, G. The effect of increasing reading rate on comprehension. *The Journal of Psychology*, 1974, *86*, 251-259.

Hornstein, S. G. A study habits program using modern learning theory principles. In F. L. Christ (Ed.), *College reading goals for the 70's.* The Western College Reading Association, 1970.

Paul, A. Diverse aspects of language development as related to reading. In M. P. Douglass (Ed.), *Reading, thought and language.* Thirty-eighth Yearbook of the Claremont Reading Conference. Claremont, Cal.: The Claremont Reading Conference, 1974.

Rankin, E. F. A simple model for the improvement of college-training programs for reading development. In P. L. Nacke (Ed.), *Diversity in mature reading theory and research.* Twenty-second Yearbook of the National Reading Conference. Boone, N. C.: The National Reading Conference, 1973.

Reid, B. S. A behavior management program for teaching. In F. L. Christ (Ed.), *Interdisciplinary aspects of reading instruction.* The Western College Reading Association, 1971.

Sherk, J. K. & Manzo, A. V. *Humanistic studies as a component of college and adult reading programs.* Paper presented at the Annual Meeting of the National Reading Conference, Houston, December 1973. (ERIC Document Reproduction No. ED 083 538)

Wilson, M. M. The reading and study center as a resource for the university academic community: A partnership in learning. In G. Kerstiens (Ed.), *Reading: Putting all the cards on the table.* The Western College Reading Association, 1972.

MICHAEL F. GRAVES
University of Minnesota

HEIDI K. HAMMOND
West St. Paul Public Schools

A VALIDATED PROCEDURE FOR TEACHING PREFIXES AND ITS EFFECT ON STUDENTS' ABILITY TO ASSIGN MEANING TO NOVEL WORDS

Teaching roots and affixes has been widely recommended (Breen, 1960; Cushenbery, 1972; Dale, 1971; Deighton, 1960; Manzo & Sherk, 1971-1972; O'Rourke, 1974; Stauffer, 1942; Thorndike, 1941). Those who advocate such teaching do so because the number of words that a developing reader must learn to deal with is extremely large and because knowledge of roots and affixes can provide students with generative tools that they can use in unlocking the meanings of thousands of novel words they encounter. Of the various word elements that might be taught, prefixes would appear to be particularly worth teaching. Briefly, the argument that prefixes are particularly worth teaching is based on three facts. First, there are relatively few prefixes, and many of them are used in a large number of words. Thus, students are not faced with an impossible number of prefixes to learn, and those that they do learn will be widely useful. Second, most prefixes have relatively constant meanings and meanings that are rather concrete and easily definable. Thus, having learned a prefix, students have a rather straightforward and easily applied meaning to work with in dealing with novel words. Third, prefixes tend to have consistent spellings. Thus, having learned a prefix, students are likely to be able to recognize it when it occurs in unknown words (Stotsky, 1976, 1977).

It therefore appears that teaching prefixes is a worthwhile practice. However, as is the case with many other teaching practices which are widely recommended and that would appear to be worthwhile, empirical results demonstrating the efficacy of teaching prefixes are lacking. In particular, detailed descriptions of validated procedures for teaching prefixes and empirical studies which demonstrate the effect of knowledge of prefixes in unlocking the meaning of new words are not available in the published literature. The present paper presents a detailed description of a specific procedure for teaching prefixes and the results of a study demonstrating the effects of this procedure on students' learning of the prefixes taught, their learning of specific words which were taught to illustrate the meaning of the prefixes, and their ability to use their newly learned knowledge of prefixes to unlock the meaning of novel words.

METHOD

Subjects

Subjects for the study were 180 seventh grade students in a middle to upper middle class suburb of St. Paul, Minnesota. Students were divided into high, middle, and low ability groups based on their scores on the

Nelson Reading Test (Nelson, 1962). Grade equivalents of the low ability group ranged from 3.9 to 7.3; those of the middle ability group ranged from 7.4 to 8.9; and those of the high ability group ranged from 9.0 to 10.5. Equal numbers of students at each of the three ability levels were randomly assigned to each of three treatment groups.

Materials

Materials included a 60 item test, instructional materials for a group taught prefixes, and instructional materials for a group taught whole words.

The test was a four-option, multiple-choice vocabulary test, which was divided into four sections. The first section contained the 9 prefixes that were taught in the study. The second section contained 9 relatively easy words, each of which included one of the prefixes. The third section contained 30 relatively difficult words that were taught to both the prefix and whole word treatment groups. And the fourth section contained 12 transfer words, difficult words which were not taught but which included the prefixes that were taught. This test was given as a pretest, a posttest, and delayed posttest.

The instructional materials for the prefix treatment group included nine transparencies, each of which was used by the participating teachers to provide direct instruction on one of the prefixes, and three vocabulary worksheets. Each of the transparencies introduced a single prefix, first showing the prefix attached to four root words, next giving the definition of the prefix, and next using each of the four prefixed root words in a sentence and then giving the definitions for each of these words. Three of these transparencies were paired with each vocabulary worksheet. The worksheets required students to (a) match the prefixes with their meanings, (b) identify the meanings of words using the prefixes, and (c) pair prefixes with root words. Answers were listed on the back of each worksheet so that students could correct their own work. This combination of a worksheet and three transparencies were used to teach three prefixes and ten words, with one of the prefixes being taught exemplified in four words and the other two each being exemplified in three words.

The instructional materials for the whole words treatment group were similar to those for the prefix treatment group except that prefixes were not taught. Three transparencies were used in this treatment. Each of them presented a sentence and a definition for each of ten words. Only two of the words on a single transparency contained the same prefix. Each of these transparencies was paired with one vocabulary worksheet. The vocabulary worksheet paired with each transparency contained the same words as the transparency. The worksheets required students to (a) identify the meaning of each word as it was used in the context of a sentence, (b) match each word with its definition in isolation, and (c) complete a cloze exercise in which each word was used. Again, answers were listed on the back of each worksheet so that students could correct their own work.

Design and Analysis

The study employed a 3 x 3 x 3 x 4 design with repeated measures on the last two factors. The variables were treatment (teaching prefixes, teaching whole words, no vocabulary instruction), ability (high, middle, low), test time (pretest, posttest, delayed posttest) and word type (prefixes taught, easy words, words taught, transfer words). The results were analyzed using the ANOVA and Newman-Keuls procedures.

Procedures

The pretest was given on one day, the treatment on the next three consecutive days, and the posttest on the next day. The delayed posttest was given three weeks later. Both experimental treatments took about 20-25 minutes each day. On each day, subjects in both experimental groups first participated in a teacher-directed, inductive session using the overhead projector. In the prefix group, the teacher presented an overhead and led students to (a) determine the meanings of the four words containing the prefix, (b) identify the prefix and its meaning, and (c) note how the prefix meaning contributed to the meanings of the four words. Then the teacher had students examine four sentences, each of which included a different word containing the prefix, and led them to identify the meanings of these words and the way in which the prefix contributed to their meanings. These procedures were followed with three overheads, each of which introduced a single prefix, before students went on to complete their worksheets. In the whole word group, the teacher presented an overhead showing the various prefixed words used in sentences, led students in determining the meaning of each word, and presented the dictionary definition of each word. For the whole word group, a single overhead included all the words for each day's lesson. In both experimental groups, following the teacher presentation the students were given the worksheets to complete. After they completed the worksheets, students checked their own answers. Students in both groups were allowed to keep their completed worksheets for study.

Both experimental treatments were designed to (a) get students actively involved in a discovery process, (b) give students massed practice on what they learned in the teacher-led discussions, (c) give students immediate feedback on their work, and (d) give students convenient materials to use for review.

Students in the control group used the time the experimental groups spent on vocabulary study doing free reading.

RESULTS AND DISCUSSION

The results of the study are best seen by considering the interactions, and the first interaction to consider is that between treatment and test time. The treatment x test time interaction was significant, $F(4,342) = 9.12$, $p < .01$. Inspection of the cell means for this interaction and the results of the Newman-Keuls test show that both the prefix and the whole word treatments produce significantly higher scores ($p < .01$) on the posttest and the delayed posttest than on the pretest, but that the scores produced by the no vocabulary treatment do not differ significantly at the three test times. The pretest, posttest, and delayed posttest scores for the prefix group were 72%, 89%, and 87%, respectively; those for the whole word group were 72%, 86%, and 85%, respectively; and those for the no vocabulary group were 70%, 73%, and 74%, respectively. Thus, not only do the prefix and whole word treatments produce significant results at the time of the immediate posttest; those results also hold up for the delayed posttest, which was given three weeks later.

At this point, it is clear that both the prefix and the whole word treatments produce greater gains than no vocabulary instruction, but there has been no evidence presented indicating that either of these two effective treatments is superior to the other. However, such evidence may be found by considering the treatment x test time x word type interaction, which was significant, $F(12, 1026) = 5.43, p < .01$, and some of the cell means for this interaction. More specifically, 12 of the 24 cell means need to be considered. On the test of words taught, the scores of the whole word group were slightly better than those of the prefix group on both the posttest (91% versus 87%) and the delayed posttest (88% versus 87%); however, results of the Newman-Keuls test indicated that niether of these differences was significant. On the other hand, on the test of prefixes taught, the scores of the prefix group were higher than those of the whole word group on the posttest (87% versus 76%) and the delayed posttest (84% versus 78%), and the Newman-Keuls test indicated that both of these differences were significant ($p < .01$). Additionally, on the test of transfer words, the scores of the prefix group were again higher than those of the whole word group on the posttest (81% versus 73%) and the delayed posttest (81% versus 72%), and again the Newman-Keuls test indicated that both differences were significant ($p < .01$). Thus, the prefix treatment, which took no longer to administer than the whole word treatment, not only resulted in students learning the words taught just as they did with the whole word treatment but also resulted in their learning more prefixes than the students in the whole word treatment and in their being able to use their knowledge of these prefixes to unlock the meaning of novel words.

One final outcome of the analysis should be noted. In no case is an interaction involving ability and treatment significant at $p \leqslant .01$. This fact, coupled with inspection of a variety of cell means, indicates that students at all ability levels can learn prefixes and can use their knowledge of prefixes to unlock the meaning of novel words. For example, high ability students produced 18% more correct responses on the prefixes and 17% more correct responses on the transfer words on the posttest than on the pretest; and low ability students produced 20% more correct responses on the prefixes and 28% more correct responses on the transfer words on the posttest than on the pretest.

The major conclusions of the study can be stated very briefly. Prefixes can be taught, and they can be taught to both higher and lower ability students. Moreover, when the procedures employed in this study are used, teaching prefixes and words appears to take no longer than merely teaching words. Finally, students can use their knowledge of the prefixes they are taught as a generative tool which will help them unlock the meaning of novel words.

REFERENCES

Breen, L. C. Vocabulary development by teaching prefixes, suffixes, and root derivatives. *Reading Teacher,* 1960, *14,* 93-97.

Cushenbery, D.C. *Remedial reading in the secondary school* West Nyack, New York: Parker, 1972.

Dale, E. Reading and vocabulary. *The Newsletter,* 1971, *36.*

Deighton, L. C. Developing vocabulary: Another look at the problem. *English Journal*, 1960, *49*, 82-88.

Manzo, A. V., & Sherk, J. K. Some generalizations and strategies for guiding vocabulary learning. *Journal of Reading Behavior*, 1971-1972, *4*, 78-89.

Nelson, M. J. *The Nelson reading test*. Boston: Houghton-Mifflin, 1962.

O'Rourke, J. *Toward a science of vocabulary development*. The Hague: Mouton Press, 1974.

Stauffer, R. G. A study of prefixes in the Thorndike list to establish a list of prefixes that should be taught in elementary school. *Journal of Educational Research*, 1942, *35*, 453-458.

Stotsky, S. L. *Towards more systematic development of children's meaning vocabulary in developmental reading programs for middle and upper grades*. Unpublished doctoral dissertation, Harvard University, 1976.

Stotsky, S. L. Teaching prefixes: Facts and fallacies. Language Arts, 1977, *54*, 887-890.

Thorndike, E. L. *The teaching of English suffixes*. New York: Columbia University, Bureau of Publications, Teachers College, 1941.

DR. JAMES L. LAFFEY
James Madison University
DONNA KELLY
Rockingham County Schools
BARBARA PERRY
James Madison University

THE EFFECT OF REPEATED READING OF TAPED LITERATURE ON READING ACHIEVEMENT

Culturally different students often come to the task of reading with sadly deficient backgrounds for understanding the complex relationship of language and printed material. Not only do they lack exposure to literature, but often their own language patterns bear little resemblance to the printed materials used in school reading programs.

Teaching these students to read becomes a laborious process. Skills are taught with heavy code emphasis, but time limitations often dictate neglect of proportionate exposure to printed material as well as practice in the use of the skills. Thus, the students are denied experiences which are essential for success in reading.

A number of techniques for listening while reading have been developed and utilized, primarily with remedial students, during the last ten years. Serving as a foundation for those techniques are the Neurological Impress Method (Heckleman, 1969), which involves simultaneous reading by the student and teacher, and Assisted Reading (Hoskisson, 1975), in which the child is read to after which he repeats back the words, phrases, and sentences. Assisted Reading is based on the contention that children learn to read in a manner similar to that in which they learn to talk. Both methods focus on the relationship between reading and language.

The relevance of the two methods to the technique used in this study can best be demonstrated by taking a closer look at the complex skill of reading. Basically, reading requires utilization of information through the mastery of a series of visual, auditory, and cognitive skills and processes. The proper meshing of these processes, along with adequate practice in using them, leads to one goal of reading instruction, which is reading fluency.

a series of visual, auditory, and cognitive skills and processes. The proper meshing of these processes, along with adequate practice of using them, leads to one goal of reading instruction, which is reading fluency.

The stages of information processing are described by LaBerge and Samuels (1974) as related to their theory of automaticity and the acquisition of skills. Visual information is transformed through a series of processing stages involving the visual, phonological, and episodic memory systems until it is finally comprehended in the semantic system. The two criteria for evaluating the processing are: (a) accuracy and (b) automaticity.

Accuracy is demonstrated through word recognition, phrasing, observation of punctuation, and the absence of hesitations, repetitions, and omissions. Accuracy requires attention, thus, of necessity, it precedes automaticity.

Automaticity is attained when attention is no longer required for the processes of decoding from the visual to the semantic systems; it occurs automatically. When this happens, the reader is free to concentrate on the units of meaning indicating that fluent reading is occurring.

The importance of automaticity is emphasized by LaBerge and Samuels (1974).

> It is this capability of automatic processing which we consider critical for the successful operation of multicomponent, complex skills such as reading. As visual words are processed through many states en route to meainingfulness, each stage is processed automatically. In addition, the transitions from stage to stage must be automatic as well. Sometimes a stage may begin processing before an earlier one finishes its processing. (p. 295)

Until decoding becomes automatic, it requires attention, and this creates a barrier to comprehension. Repeated readings aid in overcoming this barrier by providing the practice that is needed for the word recognition skills to become automatic. This, in turn, furthers the main purpose of repeated readings, which is building fluency (Samuels, 1979).

Three recent studies involving reading while listening have resulted in findings which appear to support this contention. In each of the studies, written material was taped and the students read along silently as they listened to the tape.

Schneeberg (1977) reported that her four-year study using reading while listening with educationally disadvantaged students in an inner-city school resulted in high gain scores. While expected gains in grade level ranged from 0.65 to 0.75, the students receiving the treatment showed gains of 1.0 and 1.2.

The technique was also used successfully with learning disabled children by Carbo (1978). She cited as an advantage the fact that the method requires the use of visual, aural, and tactile senses. An average gain of eight months in reading level was realized by her students after a three-month treatment period. The gains ranged from four months to fifteen months.

Repeated readings of taped literature yielded positive results in the reading achievement of a group of third grade students who were reading one to two years below grade level (Chomsky, 1976). The students seemed unable to apply their decoding skills to the printed page and the method provided practice as the students "memorized" their stories to the point that they could read them well unassisted.

The method of listening while reading requires active participation on the part of the student, yet it allows reading to be an enjoyable, rather than frustrating activity. It provides a means of bridging the gap between the decoding and comprehension levels while at the same time providing practice in the integration of skills.

In addition to the positive effect the method has on reading achievement, a significant effect on attitude toward reading was observed. Schneeberg, Carbo, and Chomsky all found the materials to be highly motivational. The students experienced success in reading, a first for many of them, and the authors noted a welcome development of self-confidence among them along with a new-found enthusiasm for reading. This was especially important because the lack of motivation is often a significant problem with remedial readers.

THE STUDY

The purpose of the study described in this paper was to determine the effect of repeated readings of taped literature on reading achievement of culturally different remedial students in the upper elementary grades.

It was hypothesized that there were no statistically significant differences between the experimental and control groups (1) vocabulary, (2) comprehension, and (3) accuracy test scores when each is taken separately.

PROCEDURES

Subjects

Twenty students from the fifth and sixth grades in a rural elementary school in Rockingham County, Virginia participated in the study. The experimental and control groups consisted of ten students each. Common characteristics of the students were that they were all reading from one and one-half years to two and one-half years below grade level and they were all participating in a Title I remedial reading supplementary program.

Analysis

The *Gilmore Oral Reading Test* was selected as the pre- and post-test measure of accuracy. Reliability for the Gilmore is .94 at third grade and .84 at sixth grade. Reading achievement was measured using the vocabulary and comprehension subtests of the *California Achievement Test*. It has a total test reliability of .87.

Treatment

The design used was a matched pairs design. The twenty subjects selected for the study were matched on the total reading scores from the *California Achievement Test*. The experimental treatment consisted of reading while listening to the tapes and the various comprehension activities. Both the experimental and control groups participated in the supplemental Title I program.

During the planning stages of the study, stories from a variety of sources were collected by the researchers according to two criteria. First, they were chosen to reflect the expressed and observed interests of the students in the experimental group. Secondly, appropriate and appealing literature was selected for the purpose of enriching the students' backgrounds.

The stories ranged in reading level from second to fifth grade. Material which was below the instructional levels of the students was taped for the initiation of the treatment period because many of the students were unable to read even relatively easy material with expression. It was felt that beginning with material which allowed for ease in decoding would maximize the importance of expression.

Taping of the stories was done by the researcher with whom the students were not familiar. Since the student heard the teacher's voice on a daily basis for routine procedures, it was decided that an unfamiliar voice would be preferable.

For their first selection, the students were offered only the stories which were at their independent levels. Since they read with poor expression and usually did not observe punctuation, the material did present a challenge. The level also allowed for realization of success fairly quickly since the students' task, essentially, was to imitate a good model.

During the treatment period of 13 weeks, the students spent approximately 15 minutes per day listening to the taped stories. After listening to the story two or three times, they were encouraged to read along orally as they listened. This was repeated until they were able to read the story fluently.

When the students mastered a story, they were given the option of reading it to the teacher, the experimental group, their classroom reading group, or another class in the school. A surprisingly large number of them chose to read it first to the experimental group and then to a different class. It is noteworthy that these students rarely read aloud willingly prior to the treatment.

A number of activities were used as a supplement to the taped stories for the purpose of promoting and evaluating comprehension as well as creating an interaction between the printed material and the use of language. These included oral discussions, writing of sequences, questions, descriptions, and comparisons, as well as creative writing.

RESULTS

Data analysis using a correlated t test revealed the following: t (9) = 1.45, $p < .20$. The means were 14.3 and 13.5 for the experimental group and control group, respectively. Thus, the hypothesis of no significant differences for vocabulary was not rejected.

There was a significant difference for the comprehension scores: t (9) = 2.81, $p < .05$. The means were 25.2 and 23.3 for the experimental group and control group, respectively. Thus, the hypothesis of no signficant differences for comprehension was rejected.

There was no significant difference for the accuracy scores: t (9) = 2.076, $p < .10$. The means were 38.6 and 34.5 for the experimental and control group, respectively. Thus, the hypothesis of no significant differences for accuracy was not rejected.

DISCUSSION

The comprehension, vocabulary, and accuracy scores were consistently higher for the experimental group than the control group. While vocabulary and accuracy scores were not statistically significant, the authors contend they were educationally significant because of extremely positive attitudes of the students toward reading after the program. Nevertheless, the interpretation of the results must be viewed with cautious optimism since a number of other variables could have affected the scores. Among them are: teacher enthusiasm for the method, the motivational aspect of the use of tape players, and the comprehension activities used. In addition, some uncertainty persists as to whether a treatment period of such short duration could have had an effect on achievement scores.

A secondary aspect of the study dealt with the observable changes in the affective behavior of the students. A collection of anecdotal records compiled on a daily basis indicated positive changes in the students' attitudes toward reading as well as changes in their self-concepts as they were perceived throughout the treatment period through informal observation.

Eight of the ten students utilized their free time two or three times during the week to return to the reading room for reading/listening sessions of fifteen to thirty minutes. Prior to the program, this did not occur. They enjoyed the stories, took pride in reading them well, and competed among themselves to read the most books.

In conclusion, the researchers believe that teachers need to be cognizant that many good readers seem to simply absorb the subskills, while remedial students may labor for years with difficulty. The good reader, because of success and the resultant love of reading, naturally pursues numerous opportunities for practicing reading skills. Practice leads to automaticity. The remedial student, on the other hand, tends to shy away from practicing reading because of lack of success and frustration — avoidance of reading becomes a habit. Thus, when some of the various subskills are eventually mastered, the remedial student continues to ignore reading as much as possible. Without practice, the experience of fluent reading is not enjoyed. The method of repeating reading while listening provides an appealing means of getting practice while furthering the goal of fluency.

REFERENCES

Buros, O. K. *The seventh mental measurements yearbook*. Highland Park, New Jersey: The Gryphon Press, 1972.

Carbo, M. Teaching reading with talking books. *The Reading Teacher*, 1978, *32*, 267-273.

Chomsky, C. After decoding, what? *Language Arts*, 1976, *53* (3), 288-296.

Heckleman, R. G. A neurological impress method of reading instruction. *Academic Therapy*, 1969, *44*, 277-282.

Hoskisson, K. The many facets of assisted reading. *Elementary English*, 1975, *53*, 312-315.

LaBerge, D. & Samuels, S. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 292-323.

Samuels, S. J. The method of repeated readings. *The Reading Teacher*, 1979, *32*, 403-408.

Schneeberg, H. Listening while reading: A four-year study. *The Reading Teacher*, 1977, *32*, 629-635.

BERNADETTE PUDIS
ELIZABETH SULZBY
Northwestern University

THE EFFECTS OF THREE INSTRUCTIONAL ACTIVITIES ON CHILDREN'S RESPONSES TO THE OPEN-ENDED QUESTION: WHAT IS A WORD?

In recent years attention has been focused upon children's development of linguistic awareness as they acquire the ability to read and to write (Downing, 1969; Ehri, 1975; Francis, 1973; Lundberg, 1978; Reid, 1966; & Vygotsky, 1962). In learning to read and to write, children must treat linguistic units as the objects of thought and action; such reflection is different from the tacit use of language which children normally engage in.

Children's awareness of the linguistic unit "word" has been one of the focal points of linguistic awareness research (Berthoud-Papandropoulou, 1978; Holden & MacGinitie, 1972; Meltzer & Herse, 1969; Myers & Paris, 1978; Papandropoulou & Sinclair, 1974; and Sulzby, 1978, 1979). One rationale for studying the child's developing concept for the linguistic unit "word" is that this unit seems to be more understandable to naive speakers than other more technical terms such as "syllable" or "sentence" (Berthoud-Papandropoulou, 1978). Further, there is great stress upon word learning in early reading instruction. Sulzby (1979b) contends that studying the child's conception for "word" has theoretical significance, for "word" is a unit in which the child confronts the dual nature of language, in which meaningless structural elements form meaningful wholes.

A number of studies (Clay, 1968; Downing & Oliver, 1974; Meltzer & Herse, 1969; Pick, 1978) have tried to identify how children come to understand the unit "word." Sulzby (1978, 1979) investigated the emergence of children's understanding of "word" as being a semantic unit as well as a unit of letters and/or sounds. She found, as have Berthoud-Papandropoulou (1978), Felzen & Anisfeld (1970), Ianco-Worral (1972), and Popandropoulou & Sinclair (1974), that, as children get older, they attend more to the semantic aspect of words. Additionally, however, she found that this increase is dependent upon how well-known the words are to the children and upon whether the words are presented in an oral or in a written mode. These findings clarify some issues related to the effects of instruction that have not been addressed in previous studies.

The Sulzby findings relate to only the most general aspects of instruction, however. It is also important to consider other more defined effects of instruction upon the development of the child's concept for "word." In a study of first graders' perceptions of printed word boundaries, Mickish (1974) found that even after a year of instruction in reading, many children were still unsure of what words were as measured by the ability to mark word boundaries. Ehri (1975) studied prereaders and beginning readers and concluded that learning to read may sharpen children's lexical awareness. Finally, Francis (1973) contended that children learn the concepts "letter," "word," and "sentence" during the process of learning to read and to write. Thus it appears that children develop their concept of "word" over time and that this development is closely related to instruction in reading and writing.

Although linguistic awareness research has pointed toward the influence of instruction, little investigation of specific details of instructional settings or models has emerged. For example, Papandropoulou & Sinclair (1974) and Berthoud-Papandropoulou (1978) speculate that instruction in reading and writing may affect children's responses to interviews about words but they do not investigate the possible effects. In a related line of research into the relationship of spoken words to real-world referents, Markman (1976) found a shift in responses when second graders were added to her sample. She attributed this shift to new knowledge coming from increased attention to phonology in school instruction. Again, no study of the effects of instruction was included.

Thus clarification is needed of the influence that instruction has upon the child's emerging concept of "word." One line of investigation of instruction would be to study children who have been taught by quite different approaches over a period of time. Another line of investigation would treat the experimental setting as being influenced by immediately preceding instructional tasks. The latter approach was taken for this study, particularly because it could be questioned whether young children were affected by artifacts of experimental settings. Sulzby (1978, 1979) used a paradigm in which children worked with words in isolation and found that young children commented primarily upon structural (letters and/or sounds) aspects of words whereas older children commented upon semantic aspects. This finding might be due to the child's perception of the experimental setting.

Based on the assumption that common instructional tasks can be simulated in experimental settings and can tap the same level of understanding that the child has for the unit "word" when doing these same tasks in the classroom, this study was designed to investigate the effects of three typical school tasks upon children's responses to a structured interview, the "What is a Word?"interview. The major focus of the study was to explore children's semantic and structural responses. It was hypothesized that, when different school tasks precede the interview, children of different ages would be affected differentially. Alternately, the tasks might have no effect upon children's responses to such an interview.

The instructional tasks were selected according to the following criteria. First, the tasks were similar to tasks frequently used in elementary classrooms. Second, the manipulation of words within each task appears to focus the child's attention upon different aspects of the unit "word." Thus a word-sorting task or flashcard activity was assumed to focus the child's attention primarily upon structural aspects of words; a directed reading activity (DRA) was assumed to focus the child's attention upon meaningful aspects of words; and a writing activity was assumed to shift the child's attention between the meaning and structural aspects.

An additional purpose of the study is to determine how the instructional tasks functioned when they immediately precede the interview. For example, they may act as priming tasks, enabling the child to access metalinguistic information, in which case different tasks might enable access of different information, particularly for the younger children. Structural tasks like flashcard activities might enable access of structural information whereas tasks focusing on meaning might enable access of semantic information.

On the other hand, the tasks may simply *teach* the child more about the concept of "word" when paired with an interview asking questions about what words are and what aspects are necessary for a unit to be called a word. In this case, we would expect the different tasks not to have different effects, but that there would be differences between an experimental group which had the three tasks paired with interviews and a control which did not have the tasks but did have the interview three times at the same interval as the experimental group.

A further possibility would be that there would be no differences between experimental and control groups, which might indicate that short-term instructional effects do not seem to be important to the child's ability to respond to a "What is a Word?" interview. This latter finding could yield some indirect evidence that children's responses to such interviews are relatively stable. This first part of the study was addressed to these hypotheses only about the first interview question, the open-ended question: "What is a word?"

METHOD

Subjects

Ninety white upper middle-class children from grades one, three, and five were chosen at random from an elementary school in a suburban school district north of Chicago, Illinois. Thirty children at each grade level were then randomly assigned either to an experimental or control group. In the school selected the method of teaching reading stressed decoding but not to the exclusion of other methods.

Materials

Instructional tasks. Each child in the experimental group took part in three instructional tasks, involving three sets of materials. For each grade level, a story of approximately 200 words was selected and served to generate materials used for each of the three school tasks. In the reading task, the story itself was read, following procedures for a Directed Reading Activity (DRA). The writing task required the child to use each twelve of preselected words from the story in a sentence that s/he wrote her/himself. The flashcard activity used the same twelve words in a word choice paradigm (Sulzby, 1978). This activity required the child to select one of two words to go with a word assigned to the child, and then to explain how those two words "go together."

"What is a Word?" Interview. A interview schedule was designed as a result of a pilot study done at the conclusion of Sulzby (1979), incorporating items from the Early Word Knowledge Test (Henderson & Gentry, 1976) and from Gibson & Levin (1975). The interview began with the open-ended question, "What is a word?" discussed herein and concluded with a number of sorting and justification tasks.

Procedures and Design

Each child was seen individually on three different occasions with approximately one week between sessions. During each session, the child assigned to the experimental group was asked to complete one of the three school tasks assigned at random in counter-balanced order. After the school task, the child was asked to respond to the interview schedule. Children in the control group did not receive the school tasks but were asked to respond to the interview schedule at the same interval. For the second and third interviews, children were told that the experimenter knew s/he had answered the questions before but that it was very important to answer them again completely, telling what s/he thought now.

The independent variables were grade level and type of school task. The dependent variable was the salience of the linguistic unit "word" with the dependent measure being the number of semantic explanations obtained during each interview session. In the following analysis, only the responses to the open-ended question, "What is a word?" were analysed. Responses were classified as semantic if the child's comments included the following information (responses were scored by two independent raters and a third rater was consulted in the case of disagreements):

Information	**Example of Response**
Words are units that are meaningful. Child must use the term *mean* in his/her response.	Words are meanings. They mean something. They have a meaning.
Words are units that enable us to communicate with each other. Child must use the term communicate in his/her response.	Something to communicate with. People wouldn't be able to communicate without words. It's communication.
Words are units that convey information: describe, tell, ask, express, understand. Terms of this type must be used in the child's response.	Something that describes something. It is something to express yourself. It could be the name of something.
A word is a group of letters that mean something.	It's some letters made up together that mean something. It's a group of letters that mean something.
A word is letters that describe, name, tell, ask, express, or help us understand.	It's a group of letters that stand for one thing or a couple of things. A bunch of letters put together to describe where or what something is doing.

RESULTS

In a preliminary analysis of the results, chi square tests were used to determine whether the frequency of semantic responses differed across grade levels, within school tasks. Chi square tests were also used to determine whether the frequency of semantic responses differed between experimental group and control group within each grade level. For this analysis, experimental group responses to the first interview session, collapsed across school task, were compared to the control group's responses to the first interview session. In addition, phi coefficients were obtained for each chi square test. Finally, Yates' correction for continuity was applied. (See Table 1.)

Table 1
Analysis of Responses To The Question: What Is a Word?

Flashcard Activity	**Chi-Square**	**Phi Coefficient**	**Yates' Correction**
Grade 1 vs. Grade 3	.370	.1111	. . . .
Grade 1 vs. Grade 5	11.62***	.6223	9.18**
Grade 3 vs. Grade 5	7.77**	.5443	6.80**
Reading Activity			
Grade 1 vs. Grade 3	3.33	.333	1.48
Grade 1 vs. Grade 5	15.00***	.707	12.15***
Grade 3 vs. Grade 5	6.65**	.4708	4.88*
Writing Activity			
Grade 1 vs. Grade 3	3.33	.333	1.48
Grade 1 vs. Grade 5	15.00***	.707	12.15***
Grade 3 vs. Grade 5	6.65**	.4708	4.88*
Experimental vs. Control Groups			
Grade 1 (E) vs. (C)	2.142	.2672	.54
Grade 3 (E) vs. (C)	.1435	.0691	. . . .
Grade 5 (E) vs. (C)	2.22	.2721	1.25

***$p < .001$
**$p < .01$
*$p < .05$

$df = 1$ in all analyses given here

The results indicate that fifth graders gave significantly more semantic responses than either third or first graders. There were no significant differences between first and third graders. The differences between fifth grades and the other two age groups were evident for all three school tasks. For purposes of this analysis, results were collapsed across school tasks and no significant differences were found between the experimental and the control group at each of the three grade levels.

DISCUSSION

In the preliminary analysis, there is evidence of an increase in semantic responses school-aged children use when the open-ended question: "What is a word?" is asked. The increase in semantic responses is consistent with the findings of previous research (Bach & Underwood, 1965; Felzen & Anisfeld, 1970; Ianco-Worrall, 1972; Papandropoulou & Sinclair, 1975; Sulzby, 1978, 1979). We found no evidence in the analysis of the open-ended question, "What is a word?" that different tasks affected the frequency of semantic responses nor that the tasks as a whole taught the experimental children information not available to the control children.

Certainly, these results are tentative and await further analysis from other parts of the interview schedule. The tasks selected may not have isolated kinds of information in the way they were designed to or it may be that short-term tasks do not affect children's responses to interview schedules of this sort. The consistent finding of an increase in semantic responses in response to abstract linguistic questioning like "What is a word?" adds support to the claim that older children find meaning more salient than structure when considering what is critical for calling a unit a "word."

What this study cannot do is address the issue of the effects of long-term instructional effects. An important next step would be to investigate whether children who are instructed consistently within different instructional models (for example, a structured phonic approach versus a whole-language approach) would exhibit different response patterns to linguistic awareness interviews and related tasks.

REFERENCES

Berthoud-Papandropoulou, I. An experimental study of children's ideas about language. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language.* New York: Springer-Verlag, 1978.

Clay, M. M. *Emergent reading behavior.* Unpublished doctoral dissertation. University of Auckland, Auckland, New Zealand, 1966.

Clay, M. M. A syntactic analysis of reading errors. *Journal of Verbal Learning and Verbal Behavior,* 1968, *7,* 434-438.

Downing, J. How children think about reading. *The Reading Teacher.* 1969, *23,* 217-230.

Downing, J. & Oliver, P. The child's conception of 'a word.' *Reading Research Quarterly,* 1973-74, *9,* 568-582.

Ehri, L. C. Word consciousness in readers and prereaders. *Journal of Educational Psychology,* 1975, *67,* 204-212.

Francis, H. Children's experience of reading and notations of units in language. *British Journal of Educational Psychology,* 1972, *83,* 551-557.

Gibson, E. J., & Levin, H. *The psychology of reading.* Cambridge, Mass.: MIT Press, 1975.

Henderson, E. H., & Gentry, R. *Early Word Knowledge Test.* Mimeograph. Charlottesville, VA: The McGuffey Reading Center, University of Virginia, 1976.

Ianco-Worrall, A. D. Bilingualism and cognitive development. *Child Development,* 1972, *43,* 1390-1400.

Lundberg, I. Aspects of linguistic awareness related to reading. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's concept of language.* New York: Springer-Verlag, 1978.

Markman, E. M. Children's difficulty with word-referent differentiation. *Child Development,* 1976, *47,* 742-749.

Meltzer, N. S. & Herse, R. The boundaries of written words as seen by first graders. *Journal of Reading Behavior,* 1969, *1,* 3-14.

Mickish, V. Children's perception of written word boundaries. *Journal of Reading Behavior,* 1974, *6,* 19-22.

Myers, M. & Paris, S. G. Children's metacognitive knowledge about reading. *Journal of Educational Psychology,* 1978, *70,* 68-690.

Papandropoulou, I. & Sinclair, H. What is a word? Experimental study of children's ideas on grammar. *Human Development,* 1974, *17,* 241-258.

Pick, A. D. Perception in the acquisition of reading. In F. B. Murray & J. J. Pikulski (Eds.), *The acquisition of reading: Cognitive, linguistic, and perceptual prerequisites.* Baltimore: University Park Press, 1978.

Reid, J. F. Learning to think about reading. *Educational Research,* 1966, *9,* 56-62.

Sulzby, E. Children's explanations of word similarities in relation to word knownness. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice.* 27th National Reading Conference Yearbook. Clemson, S.C.: National Reading Conference, 1978.

Sultzy, E. Semantic salience in relation to word knownness. In M. Kamil & A. Moe (Eds.), *Reading research: Studies and applications.* 28th National Reading Conference Yearbook, Clemson, S.C.: The National Reading Conference, 1979.

Sulzby, E. *Meaning and the child's concept of "word."* Paper presented at the NCTE Preconvention Conference, "Impact of Child Language Development Research on Curriculum and Instruction." San Francisco, November 19-20, 1979.

Tovey, D. R. Children's perception of reading. *The Reading Teacher,* 1976, *29,* 536-540.

Vygotsky, L. S. *Thought and language.* Cambridge, Mass.: The MIT Press, 1962.

◆

EARL F. RANKIN
University of Kentucky

ANN MARIE HAASE
ADELA STEWART
University of Arizona

ELISSA HOWARD
Sunnyside School District, Tucson

THE EFFECT OF A SEQUENCE STRATEGY ON TEACHING READING COMPREHENSION WITH THE CLOZE PROCEDURE

The purpose of this study is to determine the effect of one of several cloze sequence strategies for teaching comprehension postulated by Rankin (1977). The term "sequence strategies" is used to denote teaching plans useful in attaining an educational goal by organizing instruction in a given order of succession. The strategy under investigation in this study involves a three-step sequence for introducing the cloze procedure as an instructional tool. The proper introduction of the cloze procedure is important because cloze exercises tend to be novel and anxiety provoking to students. This strategy was chosen for investigation because the development of a proper state of readiness for taking a cloze exercise is basic to the success of all other strategies. If cloze materials are to be helpful in improving comprehension, a strategy which would make these materials more acceptable to students might be expected to produce superior results in contrast to simply passing out cloze exercises to be filled in without a suitable orientation.

Before describing the sequence strategy to be employed in this investigation, let us pause to consider why the cloze procedure should be expected to be a useful technique for teaching reading comprehension. Correlational evidence concerning concurrent and predictive validity of cloze tests would suggest that cloze materials might be useful for this purpose. Substantial correlations have been reported for many years between cloze tests and tests of general reading achievement, specific comprehension tests covering the same material as the cloze tests, and individual reading skills at all levels from understanding facts to getting main ideas and drawing interpretations. From a theoretical point of view, the cloze procedure is compatible with theories of communication, perception, learning, and information processing. The ability to fill in cloze blanks is an indication of language correspondence between a message source and receiver. Making correct cloze responses is compatible with concepts of perceptual and cognitive closure in gestalt theory. Cloze responses can easily be interpreted in terms of operant conditioning with stimuli in the context serving as discriminative stimuli and correct responses with feedback from the teacher as reinforcing stimuli. Rankin (1978) has viewed responses to cloze exercises in terms of basic principles in a psycholinguistic view of reading as processing information through use of semantic, syntactic, and orthographic redundancy in the context preceding and following deletions. Indeed, viewing reading as "information process-

214

ing," the cloze procedure might well be considered as a "slow motion simulation" of the reading process, although in a cloze exercise the gaps are chosen by the writer rather than the reader and otherwise available orthographic clues are missing from the deleted portions of the text. Of course this assumption is not compatible with every theoretical construct of the reading process.

One might therefore expect practice on a task which so closely simulates the reading process itself to enhance reading skill. Unfortunately, previous research on the effects of using cloze exercises for the improvement of reading comprehension has produced mixed results. Bloomer (1962) achieved positive results on grade point average for college students. Schneyer (1965) did not get positive results using 6th grade readers. Heitzman and Bloomer (1966) failed to get positive results in their study using students in grades 5 through 11. Jongsma (1971) reviewed most of the work on this topic that had been carried out up to that time and concluded that generally the teaching of reading with the cloze procedure was too mechanical and had not made use of basic principles of teaching and learning. Kennedy and Weener (1973) used visual and auditory cloze training with remedial 3rd graders to good effect. Bortnick and Lopardo (1973) made many useful suggestions for using the cloze procedure in a non-mechanical way. An excellent recent review of literature on this topic by Grant (1979) cited only a few studies which demonstrated the effectiveness of the cloze procedure for teaching reading but contained many excellent suggestions for carefully determining procedures to follow when doing cloze exercises.

Rankin (1977) suggested thirteen sequence strategies which might be helpful for the improvement of comprehension with cloze exercises. The particular strategy to be investigated in this study involves introducing the cloze procedure in a three-step sequence. First, aural cloze is used, second, aural-visual cloze is used, and third, visual is used. In an aural cloze exercise, the teacher reads orally the cloze passage, stopping at choice points where words are omitted while the students write the deleted words. In an aural-visual cloze exercise, the material is read by the teacher while the students read the printed cloze passage silently. Thus, there are two sources of clues for the student: the facial expression and the vocal inflections provided by the teacher's voice and the opportunity to study clues on either side of the missing word. The visual cloze exercise involves, as the term suggests, only the printed materials without any vocal rendition by the teacher.

This sequence assumes that filling in gaps in oral communication is an experience which people are more familiar with than filling in perceptual and cognitive gaps in printed material. Also, the use of redundant clues provided by the teacher and the printed text should make for an easy transition to filling in the missing words in a printed cloze exercise.

The general design for this study involves "sequence strategy" as an independent variable and three comprehension measures as dependent variables. Two experimental groups and a control group were used. The first group followed the previously described sequence strategy while the second experimental group used conventional visual cloze exercises. The control group followed the same reading curriculum without cloze training. Measures of comprehension include scores on a conventional standardized reading test and cloze tests scored both with exact word scoring and exact-plus-synonym scoring. The same comprehension measures were used as pre-training and post-training tests in this investigation. The study was conducted as a two factor mixed design with repeated measures on the second factor. Data were analyzed by using analysis of variance and analysis of covariance.

METHOD

Sample

Subjects consisted of 86 seventh grade students in three intact classes selected at random from six classes taught by the same instructor. Seventh grade students were chosen because of administrative convenience. Other grade levels might have been used to study this strategy. Of the total group, 29 were Anglo, 48 were Mexican-American, three were American-Indian, two were Black, and four were of other origins. The distribution of languages spoken as a first language was as follows: one spoke an Oriental language, five spoke a combination of English and Spanish, 14 spoke Spanish, and 66 spoke English. Distribution of languages currently spoken in the home was as follows: forty-nine spoke English, eight spoke Spanish, 28 spoke English-Spanish, and one spoke English-Indian. The average grade equivalent total reading score on the *Stanford Diagnostic Test* was 7.4 with a range between 4.0 and 13.0.

Materials

Students were administered the same form of the *Stanford Diagnostic Test* and a fifty deletion cloze test with every 7th word omitted both before and after six weeks of instruction. Alpha coefficient reliability on the cloze test was .74. Twenty cloze passages of between 200 and 225 words in length were used as training materials. These passages were based upon non-fiction selections and were written at a 6.5 level of readability as indicated by the Fry graph. Each passage contained 15 deletions and was constructed with every seventh word deleted. Cloze passages occurred near the end of longer selections which the students either read or heard the teacher read. This was done to provide a more adequate background of information on the topic for this sample of students with a variety of different cultural backgrounds.

Procedures

A Control Group took both pre- and post-tests and worked with the regular seventh grade reading curriculum during the six-weeks period. The Visual Cloze Group took the same tests and worked through the 20 cloze passages, receiving two passages for three days per week during the first two weeks, and one passage for two days a week during each of the following four weeks. During the first two weeks, study papers were merely handed out to be completed without feedback. During the last four-week period, study papers were completed with the exact-word correct answers given by the teacher without any discussion of the reasons for the correctness of these answers.

The Sequence Strategy Group worked on the same cloze passages administered on the same time schedule as that followed by the Visual Cloze Group. This group received the Aural Cloze treatment during the first two weeks by listening to the teacher read the selection both prior to the cloze passage and during the cloze passage itself. The teacher substituted the word "blank" for each deletion. Students wrote their words for each deletion as the exercise progressed. During the second two weeks, the Aural-Visual treatment was administered. During this phase, the same procedure was used as previously described for the Aural Cloze treatment except that the teacher first read the entire selection down to the beginning of the cloze passage and then passed out the cloze passage to the students. The students then read the cloze exercise silently and filled in the blanks as the teacher read the cloze passage orally. During the last two weeks, the Visual Cloze treatment was presented. The students read the intact passage down to the cloze passage and then completed the cloze exercise without any simultaneous oral reading by the teacher. At each step in the sequence, discussions were held covering the exact words deleted, acceptable synonyms, and the clues which offered help in making inferences for filling in the blanks.

All three groups were taught by the same teacher and used identical materials in the reading curriculum. Students in the two experimental groups were informed that their work on the cloze exercises would not influence their grades in the course.

Data Analysis

Analysis of variance results comparing Control Group, Visual Cloze Group, and Sequence Strategy Group results for the pre-tests indicated no significant differences among the three groups for the Stanford Reading Score and the Exact Cloze Score. However, on the Exact-Plus-Synonym Cloze scores, the Sequence Strategy Group had lower mean cloze scores than the other groups. Because of prior expectations that the groups might differ significantly and the fact that they were significantly different on this test scored for synonyms, an analysis of covariance was used to analyze the differences among the three groups for each dependent variable (i.e. the Exact Word Cloze measure, the Exact-Plus-Synonym Cloze measure, and the *Stanford Diagnostic Text*). The covariant for each post-test measure was its pre-test, which was an identical form of the same test.

RESULTS

Significant differences among adjusted post-test group means were found for the Exact Cloze measure and the Exact-Plus-Synonym measure (F=11.315, df=2,78, p=.001) and (F=15.942, df=2,78, p=.001) respectively. Using Tukey's HSD to determine the significance of the difference among means, it was shown for Exact Cloze scoring that the Sequence Strategy Cloze Group (M=24.03) produced significantly more exact closures than either the Visual Cloze Group (M=23.37) or the Control Group (M=21.07) and that the Visual Cloze Group produced significantly more correct closures than the Control Group.

Results of the Tukey's HSD for the Exact-Plus-Synonym scores across group means reveal that the Sequence Strategy Cloze Group (M=45.80) produced significantly higher scores than the Control Group (M=41.63) or the Visual Cloze Group (M=39.26) and the Visual Cloze Group produced significantly lower scores.

No significant differences were found among group means for the *Stanford Reading Test* results.

DISCUSSION

These results demonstrate that the use of a three-step sequence strategy proceeding from aural cloze to aural-visual cloze and finally to visual cloze produces superior results on cloze test performance whether cloze tests are scored by the exact word scoring system or by a combination of exact word plus synonym scoring. Since the visual cloze test used as the dependent variable is the same type of cloze material used by the Visual Cloze group, whereas the Sequence Strategy Group only spent the last two weeks of the six weeks period on visual cloze exercises, one might have expected the Visual Cloze Group to perform better on the post-cloze test in terms of practice effects and transfer of training principles. Since the Sequence Strategy Group performed better than both the Visual Cloze Group and the Control Group, it is recommended that this procedure be used in teaching reading comprehension with the cloze procedure. (It should be noted that this sequence strategy was used in combination with verbal feedback in the discussion of results. The writers assume that such reinforcement is basic to the effective use of any of Rankin's sequence strategies for teaching reading with the cloze procedure.) This strategy appears to compensate for the novel and sometimes threatening experience of filling out cloze exercises and may well decrease student anxiety in working through such materials where blank spaces are, in fact, unpredictable. In comparing the class atmosphere of the Sequence Strategy Group with the Visual Group, far more enthusiastic attitudes were observed with more animated discussion by the students.

These results are particularly encouraging considering the ethnic background of the students in this experiment. It is most impressive to find that cloze exercises can be helpful for English speaking students with varied ethnic and language backgrounds and that this sequence strategy works with these students for whom the cloze procedure might be particularly difficult. A subsequent analysis is planned to compare the effects of this training upon Anglo students as compared with other students who come from homes where other languages are spoken.

It is interesting to note that although this sequence strategy produced superior reading comprehension as measured by cloze tests, no effects were observed on comprehension as measured by a standardized reading test. For those who question the validity of cloze tests as measures of comprehension, such as Carroll (1952) and Tuinman, et al. (1975) this outcome may be considered a negative finding. However, students of research on readability are aware that when cloze test results are used as the dependent variable in predictive validity studies, the linguistic elements involved in multiple regression equations correlate more highly with cloze test results than any other criterion of differences in reading difficulty of materials. Standardized test results are possibly not as sensitive to the short term effects of linguistic elements upon comprehension as are cloze tests. It is the position of the authors, that the cloze test as a measure of information processing in reading (Rankin, 1978) constitutes a more sensitive test for short term treatment effects in reading comprehension than standardized reading tests. Of course, this is a theoretical assumption which does not necessarily follow from the findings of this study.

In conclusion, it is suggested that further investigations be undertaken to test the validity of other individual sequence strategies and combination of such strategies which might be useful in teaching reading comprehension with the cloze procedure. It is also recommended that the cloze procedure should be particularly useful in preparing teacher-made exercises for use in teaching students of varied ethnic backgrounds because of the ease with which materials can be matched to students with respect to difficulty level, topics of interest, and style of language.

ACKNOWLEDGEMENT

The authors wish to express their appreciation to Mr. John Luiten for his technical assistance on this project.

REFERENCES

Bloomer, R. H. The cloze procedure as a remedial reading exercise. *Journal of Developmental Reading*, 1962, *5*, 173-181.

Bortnick, R., & Lopardo, G. S. An instructional application of the cloze procedure. *Journal of Reading*, 1973, *16*, 296-300.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington: V. H. Winston & Sons, 1972.

Grant, P. L. The cloze procedure as an instructional device. *Journal of Reading*, 1979, *22*, 699-705.

Hietzman, A. J., & Bloomer, R. H. The effect of non-overt reinforced cloze procedure upon reading comprehension. *Journal of Reading*, 1967, *11*, 213-223.

Jongsma, E. The cloze procedure as a teaching technique. In R. A. Earle (Ed.), *Reading information series: Where do we go?* Newark, Delaware: International Reading Association, 1971.

Kennedy, D. K., & Weener, P. Visual and auditory training with the cloze procedure to improve reading and listening comprehension. *Reading Research Quarterly*, 1973, *8*, 524-541.

Rankin, E. F. Sequence strategies for teaching reading comprehension with the cloze procedure. In P. D. Pearson (Ed.), *Reading: Theory, research, and practice.* Twenty-sixth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1977.

Rankin, E. F. Characteristics of the cloze procedure as a research tool in the study of language. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice.* Twenty-seventh Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1978.

Schneyer, J. W. Use of the cloze procedure for improving reading comprehension. *The Reading Teacher*, 1965, *19*, 174-179.

Tuinman, J. J., Blanton, W. E., & Gray, G. A note on cloze as a measure of comprehension. *The Journal of Psychology*, 1975, *90*, 159-162.

EUNICE N. ASKOV
The Pennsylvania State University

KARLYN KAMM
ROGER W. KLUMB
University of Wisconsin-Madison

J. JACKSON BARNETTE
University of Virginia

STUDY SKILLS MASTERY: COMPARISONS BETWEEN TEACHERS AND STUDENTS ON SELECTED SKILLS

The emphasis on experimentation in the schools in the 1960's has given way to a cry of "Back to the Basics" in the 1970's. Parents are strongly questioning what their children are learning in school, particularly whether their children are learning the basic skills. The purpose of this study is to determine whether students and teachers know a particular subset of the basic skills, the study skills.

Historically, there has been some emphasis on study skills, but it seemed to decline through the 20th century until the 1950's. The renewed interest in study skills was not immediately realized by curriculum designers and textbook writers. In 1972, a survey (Otto, Kamm, & Weible) revealed that little attention had been paid to the creation of an organizational or developmental framework for the skills. Most of the skill outlines that had been devised suggested that there is a vast number of unrelated skills rather than that there are only a few fundamental skills, each with a number of related subskills. Little thought had been given in the literature to either the interrelationship existing among the skills or the related concepts. Teaching of the skills had been incidental rather than systematic. Standardized tests seemed to parallel the instructional materials in that they treated the skills as an array of unrelated tools the child should acquire. The skills appeared to be selected randomly for assessment.

As a consequence of this report in the early 1970's, an organized framework for study skills was developed as part of the *Wisconsin Design for Reading Skill Development* which presents study skills in a developmental sequence (Askov & Otto, 1971). Skills, introduced in kindergarten, are taught at each level of elementary and junior high school with increasing complexity and sophistication in application.

Because study skills have typically been taught incidentally, with the exception of the *Wisconsin Design*, one must question whether children have mastered these basic skills. One also must question whether teachers have mastered these skills, a logical prerequisite to teaching them to others. In surveying some college textbooks in reading that are used in undergraduate and graduate education, one generally finds a chapter or part of a chapter devoted to study skills. The treatment of study skills usually includes a discussion of content area reading, maps, graphs, book and library skills and reading rate. Word recognition skills, on the other hand, typically occupy several chapters, and perhaps such emphasis is warranted.

The question of teacher mastery of study skills was previously investigated and reported elsewhere (Askov, Kamm, & Klumb, 1977). The present study, however, is concerned with comparing teachers' performance with that of children and teacher education students. Because low functioning levels were found among the elementary teachers assessed, one has to wonder whether the adult groups were actually functioning at higher levels than the children and whether any differences would be found among preservice and graduate education students and elementary teachers.

PROCEDURES

Criterion-referenced assessment instruments were considered to be more appropriate for use in the context of mastery learning than norm-referenced measures since the intent of the study was to compare skill attainment among the various groups. Test norms would also be inappropriate for the groups of adults. The study skills tests that are part of the *Wisconsin Design for Reading Skill Development* (Kamm, Stewart, Van Blaricon, 1972) were chosen because they yield information about attainment of specific skills. The *Wisconsin Design* tests assess mastery of three areas of study skills—map, graph and table, and reference skills. Inspection of these skills indicates that they are not unique to the *Wisconsin Design,* but are also included in standardized norm-referenced assessments of study skills, such as the *Iowa Tests of Basic Skills* (Hieronymus & Lindquist, 1971). Although these areas do not comprise all the study skills, they do represent a sample of the skills that we expect to be taught in the elementary school.

The *Wisconsin Design*, Level G tests, corresponding to the end of the elementary and junior high school, were chosen for administration because presumably elementary teachers and future teachers should have at least mastered the skills that are expected of children at the completion of elementary or junior high school. The criterion for mastery was set at 80% in keeping with recommendations of the *Wisconsin Design* (Otto & Askov, 1974).

As described in Askov, et al. (1977), the Level G tests were administered by local reading consultants to 65 elementary school teachers in six schools located in Wisconsin, Illinois, and Virginia. All schools were just beginning implementation of the *Wisconsin Design* study skills program; the test taking was viewed as an opportunity for inservice education and teacher self-evaluation. The teachers, who were assured of their anonymity, taught various grade levels in the elementary school. Since the tests are mastery tests, time limits were not set. Most teachers took two or three hours to complete the Level G tests. At the completion of the testing, the tests were sent to the distributing company for machine scoring in the same manner that children's tests are scored.

The Level G tests were also administered to approximately 96 ninth graders in a middle-class urban school in Texas. As described in Kamm, et al. (1977), none of the children had worked with the *Wisconsin Design* study skills element, nor had they had any instruction in study skills other than the incidental instruction typically found in content area textbooks. These ninth graders were chosen as a representative sample to determine students' mastery of study skills since, according to the *Wisconsin Design,* the Level G skills are appropriate for learning by the end of elementary or junior high school. The skills should be mastered by that time in order for students to be able to use them independently in high school where there is a greater focus on learning content materials.

To determine whether undergraduate preservice teachers and graduate students (most of whom were also teachers or former teachers) had reached mastery on the sample of study skills included in Level G of the *Wisconsin Design*, the tests were mailed to reading educators located in 12 states (Colorado, Georgia, Idaho, Illinois, Montana, New Jersey, North Dakota, Ohio, Pennsylvania, Utah, Washington, and Wyoming), providing geographical representation of the United States. The sample included 77 graduate reading methods students in four colleges or universities and 432 preservice teacher education students in 12 colleges or universities. The reading educators agreed to give the tests in their teacher education classes to provide a national sampling of levels of study skills mastery of undergraduate and graduate education students. The tests were returned to the distributing company for scoring.

RESULTS

The data were analyzed separately for each of the ten study skills included in Level G. An analysis of variance was performed first to determine whether significant differences existed in the mean scores of the four groups (ninth graders, undergraduate education students, graduate education students, and elementary teachers). A Scheffé test was then performed to determine which of the groups differed significantly from which others.

Significant differences ($p<.001$) on the analyses of variance were obtained on all of the tests. On all of the tests ninth graders scored significantly lower than the other three groups. Results are presented in the following tables only for those tests on which other differences were found.

In solving problems on schedules or tables (Skill 7) the elementary teachers scored significantly lower than the undergraduate and graduate students (although significantly higher than the ninth graders). The elementary teachers also scored significantly lower than the graduate students on the use of the Readers' Guide (Skill 8). The results for these two tests are presented in Tables 1 and 2.

The question of mastery versus non-mastery was considered by the Chi square test. The data were analyzed separately for each test using the four groups to determine whether differences between the expected and observed frequencies were obtained. Because so many more ninth graders fell into the non-mastery category than would be expected by chance, the Chi square analyses were performed without the ninth grade group to determine whether differences existed between the expected and the observed in mastery or non-mastery among the three adult groups. Significant differences were obtained on determining differences on graphs by multiplication, Skill 5 (X^2=6.12, df=2, $p<.05$); use of the *Readers' Guide, Skill 8* (X^2=38.16, df=2, $p<.001$); card catalogs, Skill 9 (X^2=44.11, df=2, $p<.001$); and outlining, Skill 10 (X^2=6.43, df=2, $p<.05$). No significant differences among the adult groups were found on the other skills.

TABLE I
Standard Deviations, Means, and Scheffe´ Test for Skill 7

Group	n	M*	s^2
9th Grade	76	3.987	11.160
Undergraduate	436	10.555	8.381
Graduate	77	10.935	6.193
Teachers	35	9.029	6.676
Total	624	9.716	13.022

*Maximum Score = 15 items

SCHEFFE´

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$
$\bar{x}_1$ (3.987)	-------	-6.568*	-6.948*	-5.042*
$\bar{x}_2$ (10.555)		-------	-0.380	1.526*
$\bar{x}_3$ (10.935)			-------	1.906*
$\bar{x}_4$ (9.029)				-------

$^*p < .05$

TABLE 2
Means, Standard Deviations, and Scheffe´ Test for Skill 8

Group	n	M*	s^2
9th Grade	68	3.779	14.861
Undergraduate	432	12.789	6.719
Graduate	77	13.481	0.516
Teachers	21	11.619	7.348
Total	598	11.813	15.232

*Maximum Score = 14 items

SCHEFFE´

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$
$\bar{x}_1$ (3.779)	------	-9.010*	-9.702*	-7.840*
$\bar{x}_2$ (12.789)		------	-0.692	1.170
$\bar{x}_3$ (13.481)			------	1.862*
$\bar{x}_4$ (11.619)				------

$^*p < .05$

C, the contingency coefficient (DuBois, 1965), was also calculated following each Chi square analysis as a measure of the degree to which two nominal scales (groups of Ss and mastery/non-mastery) vary together. The highest C values, indicating the strongest association between group and mastery/non-mastery, were obtained for Skills 1 (*C*=.756), 3 (*C*=.570), 7 (*C*=.980), and 9 (*C*=.658).

DISCUSSION

Although only some of a possible array of study skills were chosen for assessment, the means of the student group—ninth graders—were surprisingly low. As reported elsewhere (Kamm, et al., 1977) the means of a twelfth grade group of students were generally higher than for ninth graders although most of the test means were below mastery level. It is heartening that the three adult groups (undergraduate education students, graduate education students, and elementary teachers) scored significantly higher than the ninth grade group on all tests. However, when the *mean* scores are considered in relation to the mastery level (80% correct) required for each test, only the graduate students attained a mean score above the mastery level for Skills 6, 8, and 10. The undergraduate students achieved a mean score above the mastery level on only Skill 8, while the means of the elementary teachers and ninth graders were consistently below mastery level.

One might argue that the difficulty level of the Level G tests is greater than that needed by some elementary teachers, particularly those in the lower grades. While it is true that not every teacher would teach such a complex skill as latitude and longitude (Skill 2), one would expect that a teacher would at least *be able to* teach difficult skills to advanced students assigned to an upper elementary grade.

Assumptions about high school and college students somehow "picking up" these basic skills during their studies are in question. The generally low levels of functioning among the adult groups—although higher than that of the ninth graders—seem to point to a serious problem. It appears that study skills are not being taught, or at least not being mastered, by the ninth grade and in some cases not by adulthood.

Focused and systematic study skill instruction in the classroom is needed to promote attainment of the basic skills. Ideally, this instruction should begin at the early elementary level and continue through the junior and senior high school so that students will have the skills to use them as independently as possible in their high school experiences.

Teacher educators also need to consider the implications of these findings for the preparation of undergraduate and graduate education students in reading. Since instruction in study skills generally falls within the province of the reading course, such instruction must be of a less cursory nature. College textbooks as well as instructional time should reflect the importance of the study skills area.

Extra efforts in teaching basic study skills, in addition to the methodology for teaching them, are called for since functioning levels appear to be so low. Inservice teacher education efforts also need to be directed to teachers' acquisition of study skills. One can no longer assume that teachers possess the study skills but do not know how to incorporate skill instruction into content studies. Indeed, how can teachers teach those skills that they have not mastered themselves? The answer is unfortunately apparent—they don't.

REFERENCES

Askov, E. N., Kamm, K., & Klumb, R. Study skill mastery among elementary school teachers. *The Reading Teacher*, 1977, *30* (5), 485-588.

Askov, E. N., & Otto, W. *The Wisconsin Design for reading skill development: Teacher's planning guide—study skills.* Minneapolis: National Computer Systems, 1971.

Dubois, P. H. *An introduction to psychological statistics.* New York: Harper and Row, 1965.

Hieronymus, A. N., & Lindquist, E. F. *Iowa Tests of Basic Skills.* Boston: Houghton Mifflin Company, 1971.

Kamm, K., Askov, E., & Klumb, R. *Study skills mastery among middle and high school students.* Paper presented at the International Reading Association, Miami, May 1977.

Kamm, K., Stewart, D., & Van Blaricon, V. L. *The Wisconsin Tests of Reading Skill Development: Study Skills,* Levels A-G. Minneapolis: National Computer Systems, 1972.

Otto, W., & Askov, E. *The Wisconsin Design for reading skill development: Rationale and guidelines.* Minneapolis: National Computer Systems, 1974.

Otto, W., Kamm, K., & Weible, E. *Wisconsin Design for reading skill development: Rationale and objectives for the study skills element.* Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

DAVID W. MOORE
University of Connecticut
JOHN E. READENCE
University of Georgia

A META-ANALYSIS OF THE EFFECT OF GRAPHIC ORGANIZERS ON LEARNING FROM TEXT

The role of cognitive readiness in learning verbal material has been emphasized by the theory and research of Ausubel (1960, 1968) and others for the past 20 years. Numerous studies have investigated the effect of advanced organizers on the reception, integration and retention of new concepts. Although derived from a different perspective, schema theorists and researchers also are emphasizing the role of cognitive readiness in acquiring meaning from text. However, only a few lines of classroom research have been conducted to help validate teaching strategies that incorporate this emphasis from basic research.

Investigations into the effect of structured overviews, or graphic organizers, on learning from text represent one of the first attempts to translate Ausubel's theory of cognitive readiness into regular classroom practice. The structured overview, developed by Barron (1969), is a graphic representation of the key vocabulary of a learning passage designed to show the relationships among new concepts and previously learned concepts. Structured overviews have since been termed "graphic organizers." These devices have been further distinguished between "graphic advance organizers" and "graphic post organizers" depending on whether they are used as a pre-reading or post-reading activity (Barron & Stone, 1974).

Graphic organizers are generally being recommended by secondary reading educators as one way to enhance learning from text. Recent textbooks by Estes and Vaughan (1978), Herber (1978), and Roe, Stoodt and Burns (1978), to mention a few, advocate graphic organizers for enhancing students' comprehension and retention of text material. As such, a careful look at the research evidence regarding this teaching strategy is warranted.

In order to integrate the research evidence on graphic organizers, the technique of meta-analysis described by Glass (1976, 1978) was employed. Meta-analysis, alternatively named integrative-analysis, results in a systematic, statistical compilation of research findings. It is an analysis of analyses. In this approach, the degree of association (i.e., measure of effect) between independent and dependent variables is obtained and then compiled across studies. Meta-analysis has been used to integrate the research on such topics as psychotherapy outcomes (Smith & Glass, 1977) and the effect of personalized instruction on learning (Kulik, Kulik & Cohen, 1979).

The purpose of this paper is to utilize the technique of meta-analysis to examine the effect of graphic organizers on learning from text. By integrating the findings of the research literature on graphic organizers, generalizations can be made regarding the research supporting their use as a teaching strategy.

PROCEDURE

The references that were consulted to identify research studies for this meta-analysis included *CIJE, Dissertation Abstracts, ERIC* and bibliographies of the research documents themselves. In order to be included in this analysis, a research study needed to meet two criteria: 1) one of the experimental treatments had to include a graphic organizer as described by Barron (1969), and 2) sufficient statistical information had to be reported so that measures of association could be recovered.

Measures of association were calculated by two methods. In the first method, t and F ratios were converted to point biserial (r_{pb}) and eta (η) values respectively, following equations specified by Cohen (1965). In the second method, with studies not reporting t or F ratios but reporting means and standard deviations, d values were calculated according to Glass (1976). These d's, in turn, were transformed to r_{pb} values (Cohen, 1977). The relative strength of the measures of association were characterized by Cohen's "conventional definitions": small ($\eta = .10$); medium ($\eta = .24$); and, large ($\eta = .37$) (Cohen, 1977, pp. 284-288).

If the control group or another treatment group produced means higher than the graphic organizer group, then the resulting measure of association was labelled negative and treated as such. Finally, r_{pb} and η values were each transformed to Fisher's Z' in order to treat them with descriptive statistics.

Separate measures of association, then, constitute the unit of analysis in this study. These measures were first computed from the available study statistics, and then combined according to various characteristics of their studies. Next, they were averaged, and then a standard error of the mean was computed according to Smith and Glass (1977). Finally, the means were visually compared so that the overall effect and the differential effects of graphic organizers on learning from text could be determined.

FINDINGS

Eighteen research studies that dealt with graphic organizers were located, and 16 met the criteria for inclusion in this study. Of the 16 studies analyzed, a total of 92 measures of association based on approximately 2098 total subjects were computed. (One study did not include its N or any degrees of freedom.)

Overall results of the combined studies indicate a small measure of association (M = .15; SD = .21; SE = .02; Md = .11) between treatment with a graphic organizer and subsequent learning from text. In fact, only about 2% of the variability in learning from text can be explained by knowing whether subjects were involved with graphic organizers or not. The rather large standard deviation indicates wide variability of the results. Additionally, one should note that minus-one standard deviation lies in the negative range (-.06). This reflects the presence of 20 effects wherein the graphic organizer group means were less than the other group means. The smaller value of the median versus the mean indicates that a few large values are inflating the mean. Therefore, the median may be the better measure of central tendency for the overall findings.

Since one average measure of effect across all studies may hide differential results of graphic organizers, comparison of effects according to selected study variables was performed. The first study variable compared was the type of treatment subjects received.

Interestingly, many studies supplemented their graphic organizer treatment with such activities as study guides, purpose-setting questions and separate group discussion. Six studies with 36 measures of association (M = .15; SE = .04; Md = .11) included supplemental activities with their graphic organizer treatment. On the other hand, eight studies with 51 measures of association (M = .12; SE = .03; Md = .09) included *only* a graphic organizer as one of its treatment conditions. Visual comparison of these means and their standard errors indicates that the supplemental activities provided in these studies produced practically no increase in the effect of graphic organizers on learning.

An additional point dealing with type of treatment should be noted. Specifically, four studies investigated the construction of graphic organizers by the readers, themselves, as a post-reading activity. These studies obtained a large effect (M = .39; Md = .38). However, since only five measures of association are included here, this relatively large finding should be interpreted with care.

A second variable considered in this meta-analysis was the length of treatment. Studies were broken into two groups: those utilizing one session only, versus those studies that extended the treatment to more than one session. The average extended treatment lasted 16 days. Seven studies with 33 measures of association used a one-session format, while eleven studies with 59 measures of association extended their sessions beyond one. (Eighteen studies are listed because some reported more than one treatment.) When treatment was limited to one session, a small measure of association (M = .14; SE = .03; Md = .09) resulted. When treatment was extended to more than one session, a similar, small measure of association (M = .15; SE = .03; Md = .11) also resulted. Visual comparison of these data indicates that the effect of graphic organizers is uniformly low both in short-term and long-term treatment conditions.

Another variable analyzed was the time of the post-test. Again, studies were broken into two separate groups. Nine studies with 41 measures of association assessed subjects' learning immediately after the treatment; ten studies with 47 measures of association gave a delayed post-test. The typical delay was 24 hours. Under the immediate post-test condition, a small measure of association (M = .20; SE = .04; Md = .13) resulted. For the delayed post-test group, a very small measure of association (M = .11; SE = .03; Md = .08) resulted.

Test content was the fourth variable analyzed in this study. This variable refers to the source of information assessed by the examiner. That is, test content based on general subject matter dealt with in class was distinguished from content based only on a passage read by study subjects. Six studies with 31 measures of association tested general subject knowledge after treatment, while 12 studies with 61 measures of association examined passage knowledge. Small measures of association again resulted from both groups, with the difference between the effect of graphic organizers on general subject knowledge (M = .19; SE = .02; Md = .10) and on passage knowledge (M = .13; SE = .03; Md = .11) being trivial.

A fifth variable explored was the type of knowledge measured in each study. Studies were grouped according to whether they assessed vocabulary or comprehension. Eight studies with 20 measures of association included a vocabulary assessment, while 13 studies with 68 measures of association measured comprehension. A medium effect of graphic organizers on vocabulary (M = .26; SE = .07; Md = .20) was noted, while a very small effect was obtained with comprehension as the criterion (M = .10; SE = .02; Md = .08). Thus, graphic organizers seem to affect vocabulary somewhat more than they affect overall comprehension.

Two other findings of interest concerning the use of graphic organizers were revealed by this meta-analysis. One study (Baker, 1977) examined the interaction of verbal ability with graphic organizer treatment. Analysis of this study revealed a large measure of association (M = .43) between graphic advance organizers and learning from text for subjects of high verbal ability. Studies by Bean (1978), Bean, Stigler, Hayes and Phillips (1978), Schwartz and Barron (1979) and Vaughan, Stillman and Sabers (1979) with university level subjects who normally possess somewhat high verbal ability also reveal substantial effects for graphic organizer treatments. These studies lend credence to Baker's findings.

A second interesting finding is that the one study (Estes, Mills, & Barron, 1969) that used graphic organizers with fiction revealed an overall effect of only .05. Such a small measure of association lends support to the notion that graphic organizers are better suited to expository writing than narrative writing.

SUMMARY AND CONCLUSIONS

Overall findings of this meta-analytic study reveal generally small effects of graphic organizers on learning from text. These findings preclude the unqualified use of graphic organizers as a strategy to enhance learning. Rather, possible tendencies toward relative advantages were noted when subjects constructed graphic organizers as a post-reading activity and also when vocabulary was the criterion variable. Supplemented treatments, differing lengths of treatment, the time of assessment, and the source of test content did not seem to alter the effect of graphic organizers. Finally, subjects with high verbal ability may benefit most from this activity, and it seems best suited for expository passages.

One area of further research with graphic organizers might be investigation into application of the encoding specificity principle (Tulving & Thomson, 1973). Basically, this phenomenon functions first as one learns new information under certain conditions, and second, as those same conditions stimulate retrieval of the information. In other words, prereading stimuli provided for the purposes of cognitive readiness may also be used later in a recall situation to retrieve that learning (Readence & Moore, 1979). The use of graphic advance organizers in combination with graphic post organizers to activate properties of the encoding specificity principle may deserve future investigation.

A final comment is in order with regard to the value of meta-analysis as a means to synthesize research evidence. Literature reviews are a popular means to accomplish this goal, but all too often they are unsystematic and subjective. Meta-analysis, on the other hand, applies systematic, analytical procedures to aggregate research findings akin to the spirit of the scientific tradition. As our research evidence accumulates and as our information-gathering methodology becomes more sophisticated, data synthesis becomes a crucial step between past and possible future research (Pillemer & Light, 1980). As such, meta-analysis can often provide investigators important insights as they attempt to design better and more powerful research.

REFERENCES

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1960, *63*, 267-272.

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart and Winston, 1968.

Baker, R. L. The effects of informational organizers on learning and term relationships in ninth grade social studies. In H. L. Herber & R. T. Vacca (Eds.), *Research in reading in the content areas: The third report.* Syracuse, New York: Syracuse University Press, 1977.

Barron, R. F. The use of vocabulary as an advance organizer. In H. L. Herber & P. L. Sanders (Eds.), *Research in reading in the content areas: First year report.* Syracuse, New York: Syracuse University Press, 1969.

Barron, R. F., & Stone, V. F. Effect of student-constructed graphic post organizers upon learning vocabulary relationships. In P. L. Nacke (Ed.), *Interaction: Research and practice for college-adult reading.* Twenty-third Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1974.

Bean, T. W. The influence of a structured overview on comprehension and oral reading miscues of selected college students. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice.* Twenty-seventh Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1978.

Bean, T. W., Stigler, M., Hayes, M., & Phillips, G. *The influence of a structured overview on university students' short and long term comprehension.* Paper presented at the National Reading Conference, St. Petersburg, November 1978.

Cohen, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology.* New York: McGraw-Hill, 1965.

Cohen, J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1977.

Estes, T. H., Mills, D. C., & Barron, R. F. Three methods of introducing students to a reading-learning task in two content subjects. In H. L. Herber & P. L. Sanders (Eds.), *Research in reading in the content areas: First year report.* Syracuse, New York: Syracuse University Press, 1969.

Estes, T. H., & Vaughan, J. L. *Reading and learning in the content classroom: Diagnostic and instructional strategies.* Boston: Allyn & Bacon, 1978.

Glass, G. V. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 1976, *5*, 3-8.

Glass, G. V. Integrating findings: The meta-analysis of research. In L. S. Schulman (Ed.), *Review of research in education* (Vol. 5). Itasca, Ill.: F. E. Peacock, 1978.

Herber, H. L. *Teaching reading in content areas* (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1978.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. A meta-analysis of outcome studies in Keller's personalized system of instruction. *American Psychologist*, 1979, *34*, 307-318.

Pillemer, D. B., & Light, R. J. Synthesizing outcomes: How to use research evidence from many studies. *Harvard Educational Review*, 1980, *50*, 176-195.

Readence, J. E., & Moore, D. Strategies for enhancing readiness and recall in content areas: The encoding specificity principle. *Reading Psychology*, 1979, *1*, 47-54.

Roe, B. D., Stoodt, B. D. & Burns, P. C. *Reading instruction in the secondary school.* Chicago: Rand McNally, 1978.

Schwartz, R. M., & Barron, R. F. *Teacher's acquisition of semantic relationships about reading instruction.* Paper presented at the National Reading Conference, San Antonio, November 1979.

Smith, M. L., & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *American Psychologist,* 1977, *32,* 752-760.

Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review,* 1973, *80,* 353-373.

Vaughan, J. L., Stillman, P. L., & Sabers, D. L. *Construction of ideational scaffolds during reading.* Tucson, Arizona: University of Arizona, 1979. (ERIC Document Reproduction Service No. ED 165 109)

VINCENT P. ORLANDO
Metropolitan State College

A COMPARISON OF NOTETAKING STRATEGIES WHILE STUDYING FROM TEXT

The question of how notetaking affects students' learning from lectures and texts has been discussed often in the literature (Peters, 1972; Orlando, 1979; Riley & Dyer, 1979). The most widely accepted model relating to notetaking and learning was developed by DiVesta and Gray (1972). They stated that notetaking serves two functions: 1) encoding which facilitates storage of material in long term memory; and 2) external storage of material for use in later study. DiVesta and Gray have argued that over-reliance on external storage interferes with processing of information in long term memory. However, Fisher and Harris (1973) indicated that external storage has a greater facilitating effect on later recall.

Whether notetaking from lectures facilitates encoding seems irrelevant given the demand characteristics of the typical three-a-week 16-week classroom setting. In this situation students must make a record of the material presented; not taking notes would be academic suicide. Indeed Pauk (1979) has stated that the goal of notetaking from lectures is to capture the lecture ideas in sequence and use them in later study. If external storage of material is of primary importance while taking notes from a lecture, the question of how notetaking affects studying from text is far less clear.

The purpose of this study was to determine if students' intentions while taking notes while studying from text (either encoding or external storage) had a significant effect on their post-studying recall performance. The performance on immediate and delayed recall tests of a group of subjects who used notes as an external storage technique was compared to a group of subjects who used notes as an encoding mechanism. The number of ideas reproduced by both groups while taking these notes was also compared.

METHOD

Design

Data for this study were collected over two sessions. During the first session all subjects were asked to read and study a passage while taking notes. One group of subjects (Notes/External Storage) was instructed to take a good set of notes detailed enough to be used for later study. The second group of subjects (Notes/Encoding) was instructed to take notes also, but to use the notes as an encoding mechanism to help them remember the information in the passage. It was hypothesized that subjects intending to use notes as an encoding mechanism would recall significantly more information on an immediate recall test than subjects intending to construct a good set of notes. It was further hypothesized that this effect would persist to a delayed test administered during the second session. Finally, it was hypothesized that subjects who attempted to construct a good set of notes would reproduce a significantly greater number of ideas than subjects who used notes as a rehearsal mechanism.

Subjects

The subjects were college freshmen, sophomores, and juniors enrolled in English literature and psychology classes at a selective midwestern liberal arts college who volunteered to participate in the study. Sixty subjects were randomly assigned to one of the two treatment groups.

Materials

The stimulus material used in the study was an 800-word passage on vitamins developed by Sagaria and DiVesta (1979). The passage was divided into ten paragraphs each of which was comprised of seven sentences containing one main idea. For each sentence an open ended factual question was constructed answerable by one-to-three words. The resulting 70 questions thus represented all the factual concepts presented in the passage.

From this pool of questions, three 20-item tests were constructed: 1) a Pretest administered before the subjects studied the passage to control for previous knowledge; 2) a Posttest administered immediately after the subjects studied the passage; and 3) a Delayed Test administered to subjects one week later. For each test two items were randomly selected from each of the ten paragraphs in the passage. This was accomplished to ensure that information was equally sampled from throughout the passage (items were not repeated across tests) and that the tests were of equal difficulty. Sagaria and DiVesta's tests for equivalence of measures were not rejected (p values in all cases were greater than 0.58).

Procedure

During the two sessions of data collection, subjects met in groups of between eight and ten. At the beginning of the first session subjects were administered the Pretest, after which the stimulus material was disseminated. Subjects were informed that upon completion of the study task they would be administered the Immediate test and also return one week later to take the Delayed test.

Both groups of subjects received oral instructions on the notetaking procedures to be used while studying. Subjects in both groups were instructed to read the passage one paragraph at a time, after which they were to stop and construct a set of notes. They were to continue with the passage in the same manner until it was completed. Subjects in the Notes/External Storage Group were instructed to construct a good set of notes that would be comprehensive enough to be used for later study without reference to the text. Subjects in the Notes/Encoding Group were instructed to take notes to help them remember the information in the passage, notes that could be used to rehearse the material presented so as to help reinforce this material in their memory. Twenty minutes were given to each group to complete their task and subjects were further instructed to review the passage and their notes if they finished before the allotted time.

Analysis

In order to assess the performance of the two groups on the immediate and delayed test a 2 x 2 analysis of variance (notetaking groups by test conditions) was computed. The differences between the two groups' Pretest and mean number of ideas reproduced in the notes were measured through the application of independent *t*-tests. Raw data were used in these analyses, level of significance was set at .05, and followup pair-wise comparisons were made using the LSD technique.

RESULTS

The results of the independent *t*-test on the Pretest Test data indicated that there was no significant difference (p=.58) between the mean number of ideas recalled for the External Storage (2.06) and Encoding (2.25) Groups. This result indicated that the groups were equally knowledgeable about the subject matter contained in the vitamins passage.

The results of the 2 x 2 analysis of variance on the immediate and delayed tests indicated significant effects for the notetaking groups (F=10.33), test conditions (F=318.81) and for the interaction between groups and tests (F=6.06). Followup pair-wise comparisons indicated a significant difference between the mean number of ideas recalled by the Encoding (14.97) and External Storage (12.46) Groups on the Immediate Test. This confirmed the hypothesis that subjects taking notes for encoding purposes would recall significantly more information than subjects attempting to construct a good set of notes. Thus, constructing a good set of notes interferes with the processing of material in long term memory. This hypothesis was reconfirmed by the comparison of the Posttest data which indicated that, once again, the Encoding Group (7.4) recalled significantly more mean ideas that the External Storage Group (4.6). The results of the *t*-test on the number of ideas recalled by both Groups indicated that the External Storage Group (52) reproduced significantly more ideas than the Encoding Group (31), confirming the hypothesis that intention can affect the quantity of notes taken.

DISCUSSION

Based on the results presented in this study it can be seen that students' intentions while taking notes from text can significantly affect the amount of information they can recall and the types of notes they take. Students whose intention was to construct a good set of notes compiled notes which were significantly more comprehensive than those students whose primary intention was to use notes as an encoding mechanism. However, these comprehensive notes came with a sacrifice. The Encoding Group recalled significantly more information than the External Storage Group both on the Immediate and Delayed Tests.

These results have implications for instructors, students and researchers. College reading instructors and students must consider the advantages and disadvantages of using notetaking either to facilitate encoding or as an external storage device. Students using notes for external storage must remember that concentrating on taking a good set of notes interferes with placing information into long term memory. This notetaking, therefore, can only be considered an initial phase of studying especially if the material being studied needs to be recalled later in a test situation. Once this type of notetaking is completed the student must engage in some type of active process, such as recitation, to encode the information for later retrieval. On the other hand, students using notetaking as an encoding-rehearsal mechanism must remember that the notes they take are not comprehensive and would be inappropriate for later study. Finally, researchers who compare the effectiveness of notetaking to other strategies of textbook studying must also consider the encoding vs. external storage variable in their research. For example, the comparative results (notetaking vs. rereading, etc.) obtained on post-study-tests may be attenuated if subjects in a notetaking group equate construction of a good set of notes with studying.

In summary, the results of this study indicate that students' intentions while taking notes can significantly affect the types of notes taken and the amount of information recalled. Students need to be aware of the pitfalls of notetaking used for either encoding or external storage, and select a type of notetaking concurrent with their needs.

REFERENCES

DiVesta, F. J., & Gray, S. Listening and notetaking. *Journal of Educational Psychology*, 1972, *63*, 8-14.

Fisher, J., & Harris, M. Effects of notetaking and review on recall. *Journal of Educational Psychology*, 1973, *65*, 321-325.

Orlando, V. P. Notetaking vs. notehaving: A comparison while studying from text. In M. L. Kamil & A. J. Moe (Eds.), *Reading research: Studies and application.* Twenty-eighth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1979.

Pauk, W. *What the college reading program should be doing.* Paper presented at the meeting of the National Reading Conference, San Antonio, Texas, December 1979.

Peters, D. L. Effects of notetaking and rate or presentation on short-term objective test performance. *Journal of Educational Psychology*, 1972, *63*, 276-280.

Riley, J. D., & Dyer, J. The effects of notetaking while reading or listening. *Reading World*, 1979, *19*, 51-58.

Sagaria, S. D., & DiVesta, F. J. Learner expectation induced by adjunct questions and the retrieval of intentional and incidental information. *Journal of Educational Psychology*, 1979, *71*, 242-249.

MARY SUE DILLINGOFSKI
KEN L. DULIN
University of Wisconsin-Madison

CHANGES IN READING-RELATED ATTITUDES DURING A ONE-SEMESTER PRE-SERVICE COURSE IN SECONDARY READING METHODS

In recent years, reading researchers have begun to examine the reading abilities, habits, and attitudes of both elementary and secondary pre-service teachers. Such research is based on the assumption that there is a high positive correlation between teachers' reading behaviors and students' reading behaviors. The results have not been encouraging. For example, Mueller (1973) surveyed pre-service teachers on their leisure-time activities and found that 40% ranked reading fourth or lower on a list of seven leisure activities. Hawkins (1967) found that nearly 20% of the pre-service teachers had read no books over a summer. Odland and Ilstrup (1963) report that 10% of pre-service teachers had read no book during a six-month period. In addition only 30% had read five or more books during that period.

In addition to reading habits, it would seem to be important to consider pre-service secondary teachers' attitudes toward teaching reading in the content area. As Vaughan (1977, p. 606) notes, "those teachers who are aware of and sympathetic to the difficulties their students may have reading their textbooks are likely to be the teachers who will learn. Resolution of the problem, then, may well begin with the teacher's attitude toward the problem." With the increase in pre-service secondary reading methods courses throughout the nation, more secondary teachers are now becoming familiar with techniques to help students read their textbooks. But what are their attitudes toward incorporating those techniques into their classrooms? Can a pre-service course in secondary reading methods affect a positive change in teachers' attitudes toward reading in the content area? Can such a course also encourage pre-service teachers to choose reading as a personal leisure-time activity?

The purpose of this study was to investigate the relationships between participation in a pre-service secondary reading methods course (which included sharing of leisure-reading books as well as the traditional secondary reading format) on changes in (a) attitudes toward teaching reading in the content area, (b) attitudes toward personal leisure-reading, and (c) number of leisure-reading books read during a given period of time.

The following three dependent variables were selected for examination:

1. Pre- and post-treatment scores on the Vaughan "Teaching Reading in the Content Area" inventory (1977);
2. Pre- and post-treatment scores on the Mikulecky Behavioral Reading Attitudes Measure (MBRAM) (1978);
3. Number of leisure-reading books read during comparable time segments during the semester.

Gain scores were computed and analyzed for all three variables. In addition to overall analyses, the dependent-variable differences between subjects coming from two teaching majors ("academic" and "non-academic" were statistically compared. Furthermore, two subgroups on the Vaughan inventory and MBRAM were also statistically compared.

METHOD

Subjects

Subjects were 48 undergraduate secondary teaching majors at the University of Wisconsin-Madison. Twenty-six were juniors, 22 seniors; 22 were males, and 26 were females.

Treatment

The pre-service one-semester course in reading methods met twice weekly for one hour sessions. The course consisted of two components. One component consisted of traditional topics covered in a secondary reading methods course (e.g., vocabulary development, comprehension, attitudes, reading in content areas, critical and creative reading, study skills, reading rate, etc.). The second component consisted of 20 minute book-sharing sessions once a week during the regular class meetings. The students and the instructor reported on, shared, and recommended leisure-reading books. All book-sharing was voluntary and oral; no written reports were required. The instructor frequently instigated book discussions and shared her own current leisure-reading books. Students were encouraged to bring copies of books to class to share with others.

RESULTS

After the data were gathered, they were subjected to statistical procedures to answer the following research questions:

1. Will there be significant differences between pre and post treatment scores for all three dependent variables for the overall group?
2. Will there be significant differences between the pre-to-post treatment gains on all three dependent variables for students with an "academic" teaching major (e.g., theatre, English, journalism, foreign language, social studies, science, elementary education) and for students with a "non-academic" teaching major (e.g., business education, art, physical education, agriculture, home economics)?

In addition, subjects were divided in terms of their scores on the Vaughan and MBRAM pretests. Gain and loss patterns were computed for subjects scoring in the higher and lower half of each pretest.

Overall, the total group increased mean scores on the Vaughan inventory from 70.34 to 75.97, a mean gain of 5.625, which is statistically significant beyond the $p < .001$ level. The total group mean scores on the MBRAM changed from 76.5 to 78.4 for a mean gain of 1.997, which is not statistically significant, since it reached only the level of the $p < .07$. The mean number of leisure-reading books read during comparable time periods changed from 2.2 to 3.1 books, a gain significant at the $p < .04$ level.

Patterns of gain and loss on all three variables were analyzed for students with "academic" teaching majors. The mean scores on the Vaughan inventory increased from 72.12 to 77.35, a mean gain of 5.23 which is significant at the $p < .007$ level. Mean gains for the MBRAM and Books-read variables were not significant at the $p < .05$ level.

Patterns of gain and loss were also analyzed for students with "non-academic" teaching majors. The mean scores on the Vaughan inventory increased from 68.33 to 74.40, a mean gain of 6.07 which is significant at the $p < .03$ level. Leisure-reading books read increased from 1.41 to 3.00, a mean gain of 1.625 which is significant at the $p < .05$ level.

Finally, patterns of gain and loss in all three variables were analyzed for students whose scores were in either the higher half or lower half on the Vaughan pretest and on the MBRAM pretest. On the Vaughan inventory, the mean pre-to-post gain of subjects in the higher half was 3.33 while the mean pre to post gain of subjects in the lower half was 8.69, a difference in mean gains of 5.35, which is significant at the $p < .05$ level. Differences in mean gains on the MBRAM were not significant at the $p < .05$ level.

CONCLUSIONS

Overall then, the following conclusions can be drawn:

1. That such a secondary reading methods program is of benefit if raising attitudes is one of the instructor's objectives;
2. That such a total program can raise attitudes toward teaching reading in the content area; and
3. That undergraduate students will do more leisure reading when encouraged to do so.

However, one must remember when interpreting these results that no separation of the two course components was attempted. The combination of the two, however, is somehow related to the reported gains on the dependent variables. In addition, the number of subjects is relatively small. Therefore, these results should be seen as exploratory and not conclusive.

REFERENCES

Hawkins, M. L. Are future teachers readers? *Reading Teacher*, 1967, *21*, 138-40.

Mikulecky, L. Mikulecky behavior reading attitude measure. In C. B. Smith, S. L. Smith, & L. Mikulecky (Eds.), *Teaching reading in secondary school content subjects: A bookthinking process*, New York: Holt, Rinehart and Winston, 1978.

Mueller, D. L. Teacher attitudes toward reading. *Journal of Reading*, 1973, *17*, 202-205.

Odland, N. and Ilstrup, T. Will reading teachers read? *Reading Teacher*, 1963, *17*, 83-87.

Vaughan, J. L. A scale to measure attitudes toward teaching reading in content classrooms. Journal of Reading, 1977, *20*, 605-609.

VICTORIA CHOU HARE
CYNTHIA A. PULLIAM
University of Illinois at Chicago Circle

COLLEGE STUDENTS' METACOGNITIVE AWARENESS OF READING BEHAVIORS

Much of the reading instruction rhetoric heard nowadays concerns the "back to basics" movement. College and university students are not exempt from the increased scrutiny applied to their reading and study skills. Unfortunately, few objective instruments exist for the purpose of tapping such students' reading and study skills. Standardized reading achievement tests such as the *Nelson-Denny* offer some insights about students' reading; however, they are not generally used for diagnostic projections. Many informal tests, e.g., oral reading miscue analyses, too, are relatively valueless insofar as more skilled readers, accustomed to reading silently, are concerned. For these reasons, the authors employed an introspection/retrospection technique to find out what good and poor readers think they do when they read.

A growing body of literature on metacognition — knowing what it is that one knows — supports the validity of using an introspection/retrospection technique. Judgments about one's own reading behaviors may provide a window on adults' reading comprehension processes, as oral reading miscues do for younger readers' comprehension processes. Flavell and Wellman (1977) suggest a possible causal chain of metamemory influencing memory behavior later in development, or at least a developing coordination between the two. It is likely that the same would apply to reading; that is, awareness of one's reading behaviors may influence one's reading performance. Brown (1978b), too, suggests that whereas there is often a surprising gap between what a child says he knows and how he performs, it may be possible that this gap narrows as a person reaches adulthood. Adulthood, therefore, may present the optimal time for a relationship between metacognitions about reading and actual reading behavior to manifest itself.

In reading, two study areas have been identified (Olshavsky, 1977). One concerns comprehension identification: readers' understanding that they do or do not understand what they are reading. A few researchers address the comprehension identification problem (Baker, 1979; Hickman, 1977; Myers & Paris, 1978). Taken together, their studies represent different perspectives on the larger comprehension identification problem. Greater sensitivity to reading-related knowledge at grade six than at grade two is demonstrated (Myers & Paris, 1978), that college students generally know when they do not understand what they are reading is suggested (Baker, 1979), and that good adult readers know they read for meaning is hinted (Hickman, 1977).

Beyond this basic self-awareness, a second study area concerns readers' awareness of behaviors they initiate when comprehension does fail. In her research, for example, Olshavsky (1977) found ". . .every time a subject elected *re-reading* as a strategy, he or she had first *stated failure to understand a clause*" (p. 671). Developmentally, judging from the research to date, students utilize a great number of strategies to help them read as they grow older (Annis & Annis, 1979; Myers & Paris, 1978; Reid, 1966; Robinson, 1965), although the nature of the strategies used has not been established. From an ability grouping standpoint, definite differences between good and poor readers' perceptions exist, regarding both number and types of strategies utilized to comprehend material. Using protocol analysis, thinking aloud while solving a problem, Olshavsky (1977) found good tenth-grade readers used more strategies than poor readers, although strategies used by both were comparable. Other research supports qualitative differences in strategy usage. Where poor readers apparently are unaware of using these strategies, good readers say they read selectively (Hickman, 1977; Robinson, 1977), hold a purpose in mind while reading (Hickman, 1977; Smith, 1967), and associate ideas read with previous experiences (Hickman, 1977; Letton, reported by Strang & Rogers, 1965). Strang and Rogers (1965) did note, however, that both high- and low-level eleventh-grade readers seemed to integrate newly-acquired ideas with their own personal development. Interestingly enough, study systems often cited in textbooks (e.g., The 3 R's for Academic Survival, OK5R, and SQ3R in Pauk's *How to Study in College*, 1974) advocate a variety of strategies for reading textual material — among them, selective prereading, adjusting reading speed, rereading, and reciting — yet, these strategies are infrequently mentioned in the metacomprehension literature.

The current picture of college students' metacognitive awareness of reading behaviors is piecemeal and tentative in nature. Existing research consists mainly of exploratory, loosely-constructed studies whose findings apply for the most part to small groups of high school students. The following study, therefore, was designed to expand the picture. Based on the available literature, then, the following expectations were generated:

1) College students will exhibit metacognitive awareness of their own reading behaviors, with regard both to the very basic awareness of the state of one's comprehension and the awareness of attendant strategies or behaviors in case of comprehension failure.

2) Mention of behaviors generally attributed to skilled reading will be associated with higher reading achievement scores, and mention of behaviors generally attributed to unskilled reading will be associated with lower reading achievement scores. Concurrently, awareness of utilizing many behaviors, no matter which behaviors, will be associated with higher reading achievement scores, and mention of fewer behaviors will be associated with lower reading achievement scores.

METHOD

Subjects

Sixty-six college undergraduates in reading methods courses were recruited. *Nelson-Denny Reading Test* raw scores were obtained for each student; the total raw score which weighted vocabulary and comprehension subtests was used in this study. Students' scores ranged from 35 to 162, exhibiting a mean of 95 and a standard deviation of 31.26. Possible scores ranged from 0 to 172.

Materials

Each student received one copy of Durkin's (1977) "Comprehension Instruction — Where Are You?" article, deemed appropriate for college reading methods class content.

Procedure

Following a brief introduction to the article, students were asked to read the article for meaning, while introspecting about how they were reading. An array of behavioral possibilities was offered to students in the form of specific instructions developed for introspective study, excerpting from Strang:

> As you read your selection, try to observe what you do as you read. For example, do your eyes move backward over the line of print? If so, why? Do they ever miss a line? Do you look at every single word? Do your eyes move faster than your mind? Do your thoughts wander? Where do they wander? Do you think of the meaning of each word or group words in phrases and thoughts?
>
> When you have finished your reading, write down immediately everything you noticed about your reading (Robinson, 1975, p. 34).

This standardized manner of jogging students' awareness was considered justifiable, for it was expected that students who read differently *would* retrospect differently about their reading.

RESULTS

All retrospections mentioned by students were collected and independently categorized by two judges, following behavorial categories designated in an earlier pilot study as guidelines. Behaviors were included regardless of whether they were considered to be educationally sound or unsound. Interrater reliability of .98 indicated a high degree of agreement between the judges in categorizing the nine most frequently-mentioned behaviors.

Means and standard deviations for these behaviors are presented in Table 1. Becoming Distracted, which included comments of attention-wandering

Table 1
Means and Standard Deviation of Metacognitive Reading Behaviors

Behavior	Mean	Standard Deviation
Becoming distracted	.69	.46
Rereading	.65	.48
Reading for meaning/understanding	.38	.49
Tying reading to personal experience	.29	.46
Reading every word	.27	.45
Selectively reading	.28	.43
Eyes moving faster than mind	.21	.41
Adjusting reading speed	.20	.40
Chunking words	.18	.38

Note: n = 66.

or daydreaming, was mentioned more often than any other behavior; over two-thirds of the students named this category. Rereading, mentioned by 65 percent of the students, was next in frequency. Students who wrote that they were Reading for Meaning, understanding or not understanding the text — they hypothesized a basic form of self-awareness — represented 38 percent of the total group.

Multiple regression analyses were performed to fit an equation which would explain the *Nelson-Denny Reading Test* score received by a student. Originally, the nine frequently-mentioned behaviors were entered into the analyses as potentially explanatory variables. And, to check whether the number of behaviors itself correlated with reading score, a Behavior variable was attached to the list of nine independent variables. Preliminary regressions showed certain variables to be nonsignificant predictors of variation in students' scores. These variables were eliminated, and eventually, the set of independent variables was reduced to five: Reading for Meaning, Rereading, Selectively Reading, Adjusting Reading Speed, and Behavior. In its final form, the regression equation which was able to explain 59.82 percent of the variation in the raw score appeared as follows:

Score = 61.3 + 17.8 Meaning + 14.8 Reread + 12.0 Select + 19.1 Speed + 5.6 Behavior
(12.56) (2.67)*** (2.00)** (1.51)* (2.33)** (1.35)*

Because these retrospections were expected to be associated with higher raw scores, one-tailed t-tests were appropriate for determining significance. Numbers in parentheses represent t values obtained using this regression equation. One asterisk indicates significance at the 10 percent level, two asterisks at the 5 percent level, and three asterisks at the 1 percent level.

A couple examples may help to illustrate the use of this equation. If a student mentioned none of the nine frequently-mentioned behaviors whatsoever, then that student's reading score may be estimated to be 61.3. If a student mentioned both Reading for Meaning and Adjusting Reading Speed, then the student's estimated reading score may be:

$$102.3 = 61.3 + 17.8 + 12.0 + 5.6 + 5.6,$$

based on this equation. Values of 5.6 points increased this student's score for each behavior mentioned. This is especially noteworthy in view of the fact that all nine behaviors, regardless of educational desirability, were counted in the Behavior variable.

A second regression equation was formulated which explained raw score on the basis of specific behaviors alone, eliminating the quantitative Behavior variable which was significant at the 10 percent level. Four behaviors explained 58.6 percent of the variance, as demonstrated in the following equation:

Score = 62.8 + 23.3 Meaning + 21.7 Reread + 18.4 Select + 25.0 Speed
(13.15) (4.37) (3.95) (2.87) (3.60)

Additional stepwise regression analyses demonstrated that Adjusting Reading Speed, Reading for Meaning, Rereading, and Selectively Reading retrospections contributed to 25.5, 16.1 , 11.4, and 5.6 percent, respectively, of the variation in reading score. All t values are significant at the .01 level in this equation. When the 5.6 points for mention of any particular behavior is added to each behavior in the previous equation, the effects of both equations can be seen to be similar.

DISCUSSION

College students in this study were able to report how they believed they read. In fact, students' retrospections about their reading behaviors are acute enough to discriminate significantly between high and low scores on a reading achievement test. This evidence strengthens the possibility of a causal chain between metacognitions about reading behaviors and reading behavior itself.

The data suggest that high-scoring readers are more actively involved in reading; four active behaviors — Reading for Meaning, Rereading, Selective Reading, and Adjusting Reading Speed — discriminate between high and low reading scores. Also, higher scorers are conscious of more behaviors during reading than lower scorers. Interestingly, behaviors which are marks of relative passivity in reading, that is, Becoming Distracted and Eyes Moving Faster Than Mind, are neither positively nor negatively associated with reading scores. These behaviors, although they seem like they would be decidedly detrimental to good reading, appear instead to be naturally-occurring phenomena for all readers.

Why readers retrospect differently is not yet known. As with all metacognitive data, there exists the chance that readers who score higher may only be more articulate in expressing how all readers do, in fact, read. Seeking discrepancies between how poor readers think they themselves read and how they think good readers read may help resolve this issue. If differences are perceived, perhaps intervention studies which capitalize on knowledge of good adult reading strategies may be attempted. At any rate, this study suggests an extremely practical, ecologically valid way of getting at students' reading processes.

ACKNOWLEDGEMENTS

Acknowledgements to Dr. Walter J. Wadycki, Department of Quantitative Methods, University of Illinois at Chicago Circle, for statistical help.

REFERENCES

Annis, L., & Annis, D. *A normative study of students' preferred study techniques.* Paper presented at the meeting of the American Educational Research Association, San Francisco, April 1979.

Baker, L. *Do I understand or do I not understand: That is the question* (Reading Education Report No. 10). Urbana-Champaign: University of Illinois, Center for the Study of Reading, July 1979.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glasser (Ed.), *Advances in instructional psychology.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978. (a)

Brown, A. L. Metacognitive development and reading. In R. J. Spiro, B. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978. (b)

Durkin, D. *Comprehension instruction — where are you?* (Reading Education Report No. 1). Urbana-Champaign: University of Illinois, Center for the Study of Reading, October 1977.

Hickman, J. What do fluent readers do? *Theory into Practice,* 1977, *16,* 372-375.

Myers, M. II, & Paris, S. G. Children's metacognitive knowledge about reading. *Journal of Educational Psychology,* 1978, *70,* 680-690.

Olshavsky, J. E. Reading as problem solving: An investigation of strategies. *Reading Research Quarterly*, 1977, *12*, 654-674.

Pauk, W. *How to study in college* (2nd ed.). Boston: Houghton-Mifflin Company, 1974.

Reid, J. Learning to think about reading. *Educational Research*, 1966, *9*, 56-62.

Robinson, H. A. Reading skills employed in solving social studies problems. *The Reading Teacher*, 1965, *18*, 263-269.

Robinson, H. A. *Teaching reading and study strategies: The content areas*. Boston: Allyn and Bacon, Inc., 1975.

Smith, H. K. The responses of good and poor readers when asked to read for different purposes. *Reading research Quarterly*, 1967, *3*, 53-83.

Strang, R., & Rogers, C. How do students read a short story? *English Journal*, 1965, *54*, 819-823, 829.

WILLIAM METHENY

THE INFLUENCE OF GRADE AND PUPIL ABILITY LEVELS ON TEACHERS' CONCEPTIONS OF READING

Reading theoreticians and researchers continue to focus on the nature of teachers' thought processes for reading instruction. Some maintain that teachers have implicit reading beliefs based on reading models that guide them to instruct in particular ways (Harste & Burke, 1976). A teacher with a phonics belief system, for example, would be expected to use a phonics instructional approach with readers.

Recent classroom research suggests that teachers' reading beliefs are not based on particular reading models per se, but are contingent on such contextual variables as the pupil and his/her environment (Bawden, Buike, & Duffy, 1979). A survey of some 200 Michigan classroom teachers indicated that teachers may shift their reading beliefs from a skills orientation in the lower elementary grades (k-3) to a more pupil-centered, interest-based reading approach in the upper elementary grades (4-6) (Metheny, 1979).

The influence of teachers' beliefs about learners may have an impact on good and poor readers in the kinds of instruction they receive and the kinds of reading information they use (see Allington & Chura, 1979, for a discussion of good and poor readers). Given the relationships between teachers' reading beliefs and their instructional methods (Bawden, Buike, & Duffy, Note 1), and the impact of these methods on learners' reading strategies (Barr, 1975; Cohen, 1975; DeLawter, 1975), careful consideration should be given to teachers' reading beliefs for different groups of learners. For example, if some teachers believe that low-ability readers should receive a grapho-phonic approach while the good readers need semantic-contextual experiences, these beliefs many influence their instructional methods and the learners' reading strategies accordingly. The present investigation seeks to understand how teachers think about reading with respect to their views of different kinds of readers.

The hypothesis generated from these studies—that teachers' conceptions of reading vary with grade and pupil reading—ability level—formed the basis for the present study. Teachers' conceptions of reading are defined as those principles or beliefs that govern teachers' thought processes in reading instruction. This study examines only teacher beliefs, not actual instructional methods or student reading strategies. It investigated if a group of primary teachers' reading beliefs for their current classrooms differed from those they might have if (1) they were teaching an upper-elementary grade and/or if (2) they were teaching a classroom of students from a socioeconomic background different from that of their present students. The statements each teacher made about his/her present lowest reading group were compared to those s/he made for her top reading group, for the classroom as a whole, for an upper-grade classroom, and for a different socioeconomically based classroom.

METHODS

Nine teachers (eight first-grade, one second-grade) participating in the Conceptions of Reading Project research were interviewed in May, 1979. These teachers were known to group their students for reading instruction. Each teacher was asked to state her reading beliefs for five groups of readers using the decision criteria taken from the Reading Propositional Inventory (Duffy & Metheny, 1979). These eight criteria, substantiated from the classroom observations as relevant to the kinds of decisions teachers make in reading instruction, are as follows: (1) criteria for judging pupil reading success, (2) criteria for selecting materials, (3) criteria for forming instructional groups, (4) time allocation to reading activities, (5) time allocation to pupils, (6) favored word recognition prompts, (7) comprehension emphasis, and (8) favored instructional role.

The interviews were structured in the sense that the teachers were given the criteria from which to generate their beliefs for each group. They were allowed to make as many statements as they wanted to convey these beliefs. First, the teacher gave statements for her lowest reading group, secondly, for a classroom of a different socioeconomic background, and thirdly, for her highest reading group. Then the teachers gave statements for an upper elementary-grade classroom (4th-5th) and finally, for their current classroom as a whole.

On the socioeconomic background conditions, the teachers were given situations to respond to in which the students' socioeconomic background levels were nearly the opposite of those of their current students. Because students in six of the nine classrooms were of a low to low-middle socioeconomic status (SES), their teachers were given a high SES situation to respond to. Three teachers received a low SES situation. The six teachers responding to the high SES situation were told to imagine they were teaching at their current grade level in a high SES classroom. They were told that their imaginary students came from affluent homes and had parents who were predominantly white-collar professionals such as doctors and lawyers. The three teachers responding to the low SES situation received similar instructions but were told that their imaginary students came from low income homes and had parents who were either unemployed or had low paying jobs. For the upper grade-level situation, the teachers were asked to imagine they were teaching a fourth- or fifth-grade classroom. Each interview lasted approximately one hour. The interviews were audiotaped and transcribed verbatim.

ANALYSIS AND FINDINGS

The interviews were analyzed for the words, phrases, and sentences the teachers used to describe each group of readers. These descriptions provided data for analysis.

Each teacher's statements about her lowest reading groups were compared to those she made for a contrasting SES classroom situation, her top readers, an upper-grade classroom, and her present classroom as a whole on each of the eight reading criteria. A teacher's statements for a particular group were designated as "different" if the teacher did not include any of the descriptions she used in talking about her lowest reading group on a criterion. The lowest reading group served as the comparison because it provided the most contrast in the teachers' reading beliefs across the groups.

The number of teachers whose reading conceptions for any group differed from those they held for their lowest reading group was tallied for each group on criterion. From this tally, the average percentage of teachers whose conceptions differed for each group was computed across the eight dimensions.

Some major results emerged. In descending magnitude, the statements the teachers made about the upper-grade readers, top-group readers, and high-SES readers differed the most dramatically from those beliefs they gave about their lowest group of readers. Of the nine teachers, 83% held different reading beliefs for the upper-grade and low-group readers. Similarly, 77% of these teachers had different reading beliefs about their top readers. Also, 73% of the six teachers responding to the high SES situation held dissimilar beliefs for their low-group readers and readers from high SES backgrounds. In contrast, only 17% of the three teachers stating their beliefs about low SES readers held different beliefs for them and their low group readers. Views for their present whole classrooms that differed from their statements made about their lowest group of readers were held by 36% of the nine teachers.

This analysis lends support to the hypothesis that teachers' reading beliefs vary across particular reader subpopulations. The teachers believed that upper-grade readers were conceptually the most dissimilar to their lowest group of readers, followed closely by their top group and high-SES readers. The teachers tended to share similar reading beliefs for the low-SES pupils and the low-group readers.

Further data analysis provided information on the nature of the reading belief distinctions the teachers held for the reader subpopulations. Their beliefs on each criterion for each group were rank ordered from the most to the least repeated. These ranked statements were considered the teachers' conceptions of reading for the group on a criterion. The criteria that provided the most contrast in their reading conceptions across the groups are discussed below in order of decreasing contrast.

The kinds of comprehension the teachers believed the groups should receive varied with group reading-ability level. The teachers generally believed that their lowest group and the low-SES readers should receive simple, factual, low-level comprehension exercises, while the high-SES top-group, and upper-grade readers should experience higher levels of comprehension and inferential, critical, and analytical thought.

This coincided with their beliefs for allocating time to reading-group activities. The teachers believed that rather than spend their time on oral reading and word-attack activities like the low-group and low-SES readers, the top-groups, upper-grade, and high-SES pupils should concentrate their time on high-level comprehension activities.

Not only did the teachers believe the low-group readers and low-SES students should receive simple and controlled reading experiences, most of them believed these groups needed more direct instructional time and supervision than the other groups. For example, the teachers felt that their instructional role for the low-group and low-SES readers should be that of a controlling, step-by-step director. In contrast, this instructional role shifted from that of a controlling director to a more passive, guider-facilitator for the high-SES, top, and upper-grade readers. The teachers explained in their statements for their present classroom, that they gave direction to those who needed it most; in this case, they gave it to the low-ability readers.

The relationship between reader ability levels and exposure to progressively more complex reading experiences became more evident after looking at the criteria the teachers believed they should use in selecting materials for the groups and for helping them recognize problematic words. The teachers stated that their present highest group of readers, upper-grade readers, and, to an extent, the high-SES pupils needed a wide variety of interesting materials to challenge them and enrich their reading experiences. They believed their low-group and low-SES readers needed interesting and fun materials too, but for different reasons; these students needed the materials not as a challenge but as a motivator for learning basic word-attack skills. Interesting materials were seen as devices to circumvent their reading frustrations.

The teachers believed that the low-group and the low-SES readers needed and should be given phonetic cues for recognizing unknown words while the high-ability, high-SES, top-group, and upper-grade students were seen as more equipped to use both phonetic and contextual cues from the teacher. Apparent in these statements is the theme that readers must acquire the basic reading skills before they can progress. These criteria are crucial for grouped reading instruction.

The teachers believed that their pupils essentially grouped themselves on the basis of their word-attack skills and on their learning styles. As mentioned earlier, the low groups and the low-SES readers were believed to have fewer of these skills and to be less independent and attentive workers than the high-SES, upper-grade, and top-group readers. They need more supervision, more direct instructional time, and more instructional group time (i.e., less individualization).

DISCUSSION

Several common themes emerged from this investigation. The results supported the hypothesis that teachers' reading conceptions vary for different groups of readers. Generally, the teachers held very different conceptions of readers in low groups and low-SES situations from readers in the top reading groups, upper grades, and high-SES situations. The teachers tended to equate low-SES pupils with low-ability readers and high-SES pupils with high-ability and upper-grade readers.

The teachers tended to share a view of reading that made student acquisition of certain basic word-attack skills a prerequisite to more complex reading and thinking experiences. Such experiences were reserved for those readers who had essentially learned to decode. This belief is further evident in that the teachers saw many similarities between their top primary-grade readers and the upper-grade readers. At some level, they apparently believe the two groups are separated only by maturity and progressively more difficult reading encounters.

This teacher view of reading and learners raises an important point. First, it seems the teacher's beliefs are consistent with the research findings that poor readers process reading information at the phonemic level, while good readers depend more on semantic information (Mosenthal, Walmsley, & Allington, in press). This suggests that the teachers make such a distinction between their good and poor readers, and may instruct them at these levels. The degree to which a learner depends on a particular level of reading information processing may have an important bearing on his/her status as a good or poor reader in the teachers' eyes.

It was the purpose of this investigation to understand how teachers think about their reading. Apparently teachers do not think in terms of abstract, theoretical reading models: their reading beliefs are a function of their expectations for different kinds of readers. This distinction is important in that teachers are not operating at the theoretical levels found in reading models as they face practical classroom concerns that ultimately determine how they will make their reading beliefs operational. Reading theory and research must take these factors into account to make its products relevant to classroom teachers.

Naturally, this study has limitations. A larger sample of teachers may have produced greater heterogeneity in teacher belief patterns. Since the teachers only stated their beliefs, their actual practices may differ, especially in the two hypothetical situations. No information was used concerning their past teaching experiences.

IMPLICATIONS

It appears that the reading conceptions of some teachers depend upon the group of readers being instructed. Teachers' conceptions and expectations of learners apparently interact with, and may be as important as, their conceptions of the subject matter. The nature of the learner is a contingency not to be ignored in classroom reading research on teachers' thoughts and behaviors.

These findings have a bearing on the current research on good and poor readers. The teachers in this study provided different reading experiences for the two groups. Their assumption equating reading ability levels with students' socioeconomic backgrounds may have important implications for these readers in classroom reading instruction in such areas as grouping, direct instructional time, and exposure to high-level thinking and reading activities.

ACKNOWLEDGEMENTS

The author would like to thank Gerry Duffy and the Conceptions of Reading Project staff for their contributions in data collection and manuscript review.

REFERENCES

Allington, R., & Chura, S. Effects of instruction type on word prediction behaviors. In M. L. Kamil & A. J. Moe (Eds.), *Reading research: Studies and applications.* Twenty-eighth Yearbook of the National Reading Conference. Clemson, S. C.: National Reading Conference, 1979.

Barr, R. The influence of instructional conditions on word recognition errors. *Reading Research Quarterly*, 1971, *7*, 509-529.

Bawden, R., Buike, S., & Duffy, G. *Teachers' conceptions of reading* (Res. Ser. No. 47). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

Cohen, A. D. Oral reading errors of first grade children taught by a code-emphasis approach. *Reading Research Quarterly*, 1975, *10*, 616-650.

DeLawter, J. The relationship of beginning reading instruction and visual patterns. In W. D. Page (Ed.), *Help for the reading teacher: New directions in research*. Urbana, Illinois: National Council of Teachers of English, 1975.

Duffy, G., & Metheny, W. The development of an instrument to measure teacher beliefs about reading. In M. L. Kamil & A. J. Moe (Eds.), *Reading research: Studies and applications*. Twenty-eighth Yearbook of the National Reading Conference. Clemson, S. C.: National Reading Conference, 1979.

Harste, J. C., & Burke, C. L. A new hypothesis for reading teacher education research: Both the teaching and learning of reading are theoretically based. In P. David Pearson (Ed.), *Reading: Theory, research and practice*. Twenty-fifth Yearbook of the National Reading Conference. Clemson, S. C.: National Reading Conference, 1977.

Metheny, W. *The relationship between teacher demographic variables and their reading beliefs: Three key variables*. Paper presented at the Annual Meeting of the International Reading Association, Atlanta, May 1979.

Mosenthal, P., Walmsley, S., & Allington, R. Word recognition: Toward a multicontext model. *Visible Language*, in press.

RUTH S. MEYERS
Brooklyn College

LENORE RINGLER
New York University

TEACHER INTERNS' CONCEPTUALIZATION OF READING THEORY AND PRACTICE

Research in teachers' concepts of the reading process has indicated that teachers use their internalized concept of the reading process as a guide to instruction. The work of Harste and Burke (1977) and the recent work of Mitchell (1978) support this thesis. Harste and Burke examined three theoretical orientations toward reading. Their research indicated that both teachers and learners have distinctive and identifiable internal models of reading and that their interpretations of these models affect teacher-pupil interactions significantly during reading instruction. Mitchell (1978) examined the practices of teachers as they respond to the oral reading errors of students. Her analyses suggested that the teachers' practices were related to the model of reading and learning emphasized during their training. However, there is no clear indication that teachers are aware of their belief systems or if their stated beliefs are related to their practice. Thus, teachers fail to appreciate the potency or origin of their concepts of reading. These concepts have evolved through earlier training, from imposed curriculum demands, or through the use of particular reading tests or programs.

As teacher educators responsible for training student interns in the teaching of reading, we have developed an interactive model of teacher training (Ringler, Smith-Burke, & Meyers, 1977, 1978, 1979). In this model knowledge of reading and teaching/learning are viewed as interactive processes. That is, graduate students learn to employ a model of reading in their own teaching as they participate in a supervised clinical experience. The model of reading imposed is also interactive in that reading is considered a process which involves active construction of meaning in which readers predict and confirm text (Adams & Collins, 1977; Anderson, 1977; Meyers, 1978; Rumelhart, 1976). In this model of teacher training, the role of the teacher educator is that of a facilitator who creates a context which enables student interns to a) understand and verbalize their concepts of the reading process; b) discuss the implications of these models for practice; c) acknowledge the limitations of their models which we have found tend to be linear ("top-down" and "bottom-up"); d) develop an interactive perspective of both reading and learning; and e) translate this interactive perspective to practice.

In our work with student interns we have become concerned as to whether their theoretical concepts are indeed reflected in their practice. Observation of interns, as they work with learners exhibiting reading problems, has indicated that unless their theoretical rationale is related to their practice, their choices of instructional strategies and materials tend to be inappropriate and not well targeted to learners' needs. Therefore, this study was undertaken to examine the following research question: Are student interns' verbalized concepts of the reading process consistent with their practice as reflected in their analysis of diagnostic case material?

METHOD

Subjects

The study involved 60 teacher interns in two urban universities during the second half of their masters degree program. All the participants have a background in psychology and have completed graduate course work in reading theory, educational psychology, and linguistics or psycholinguistics, and are enrolled in a year-long course in diagnosis and intervention of reading problems.

Procedure

Interns were asked to write a brief response to the following set of questions:

What is your concept of reading?
How does a reader read?
What is the reading process?

Following this, interns were presented with a case study of an 11 year old sixth grader. The case study data were designed to be ambiguous so that various diagnostic hypotheses and instructional intervention strategies were applicable (Ringler, Smith-Burke, & Meyers, 1979). That is, it would be possible to interpret the data from several theoretical frameworks, because the learner exhibited both word identification problems and reading comprehension problems and had particular strength in listening comprehension. Miscues at the word level (words in isolation and words in context) revealed that the learner had some knowledge of phoneme/grapheme correspondences and some ability to use contextual information to identify unfamiliar words. The case study, as presented to the teacher interns included a summary of standardized and informal test data, a third grade selection from an informal reading inventory scored both for oral reading miscues and comprehension, a scored retelling of a third grade narrative, and a portion of the same story coded during oral reading. In addition, interns were provided with a summary of miscues of isolated words and words in context read by the student. Interns were asked to study the case study materials and then list in order of importance their diagnostic hypotheses relevant to the student's reading performance. Following this, they were asked to list in priority order the intervention strategies that they would recommend for this reader.

This data, as well as background information on the interns' teaching experience, were collected in the sixth week of the course.

TREATMENT OF DATA

The interns' written responses of their concepts of the reading process were categorized as reflecting one of three current models of the reading process. Their model was considered top-down if they emphasized a wholistic, meaning-based concept of reading. The following statement is illustrative of those responses that were categorized as "top-down": "Reading is a process in which the reader interacts with the written text in an effort to extract information, understand concepts or underlying motives or meanings. The ability of the reader to select and understand the information he/she wishes to obtain is dependent on the skill, experiences of the reader." A bottom-up model emphasized sequential decoding of smaller to larger linguistic units with meaning treated as secondary. An example of an intern's response in this category is: "Reading takes place when understanding of the squiggles has occurred. A reader can obtain meaning from exposure to the system of print which corresponds to his spoken language." An interactive model emphasized flexible use of meaning and decoding strategies. The following response is representative of those statements categorized as interactive: "Reading is a process by which the reader brings to the text his or her own experience in order to comprehend the written word. The ability to decode is only a part of the reading process. Taking all the words and making sense of it (comprehending) is most important."

Diagnostic analyses and recommendations for interventions were similarly categorized. That is, for a protocol to be categorized as top-down, diagnosis included an assessment of prior knowledge and language of the reader, and a plan of instruction based on helping the reader utilize this background knowledge to make predictions while reading. Protocols considered bottom-up included a diagnosis which focused on an assessment of discrete skills and instruction geared to remediating these isolated skills beginning with word identification skills. For a protocol to be considered interactive, diagnosis included an analysis of the use of the reader's knowledge of phonetic, syntactic, and semantic systems as meaning was constructed from the text. Intervention included helping the reader integrate new information from the text with knowledge previously acquired. Concept statements and protocols were categorized by two raters with interrater reliability of .90. A test for consistency was calculated between the stated concepts of the reading process and the diagnostic/intervention protocols.

RESULTS AND DISCUSSION

Results indicated that there was no significant relation between teacher interns' stated concepts of the reading process and their practice as reflected in their analysis of the protocol materials. However, an examination of the crossbreak (Table 1) suggests some interesting trends.

TABLE I
Crossbreak of
Concepts of Reading Process
and Case Study Analysis

Reading Models	Case Study Analysis			
	Top-Down	Bottom-up	Interactive	
Top-Down		^{8}(.80)	^{2}(.20)	10
Bottom-Up	^{1}(.05)	^{13}(.65)	^{6}(.30)	20
Interactive	^{1}(.03)	^{19}(.63)	^{10}(.33)	30
	2	40	18	60

Of the 60 participants in the study 30 (50%) verbalized an interactive model of the reading process. Of these 30, one-third (10) showed consistency in their analysis of the protocol in that they emphasized an interactive approach to diagnosis and intervention. It is important to note that of these 10 teacher interns, nine who presented an interactive model were naive teachers in that they lacked teaching experience. The one remaining teacher intern had used an eclectic approach to teaching reading in her previous experience.

Further examination reveals that 40 of the interns indicated a "bottom-up" approach to diagnosis and teaching even though 19 of these interns verbalized an interactive approach. Those interns who verbalized an interactive approach would seem to be influenced by current theories of the reading process but have yet to apply theory to practice. The majority of this group of 40 were experienced teachers. The differences in responses between these experienced teachers and inexperienced teachers who were consistent in their use of an interactive model, deserves serious consideration. From an information-processing point of view it appears that experienced teachers bring to their college courses an array of structured strategies and processes that reinforce a "bottom-up" model. We assume that current emphasis on a diagnostic-prescriptive model of teaching with related published materials has had great influence on these experienced teachers.

Of further interest are the ten teachers who verbalized a "top-down" concept of the reading process, though none stated diagnostic hypotheses or intervention strategies reflecting this model. This is of interest because the theoretical positions of both Goodman (1976) and Smith (1978) are examined in detail and miscue analysis (Goodman & Burke, 1971) is used as a training technique during course work. The inference here is that the "top-down" model of reading does not provide teachers with what they perceive as adequate instructional strategies.

This limited study which suggests the tenacity of teachers' belief systems has several implications for teacher education programs. As an initial step it would seem to be important to elicit and examine the prior knowledge of teacher interns as it relates to teaching reading. Establishing this conscious awareness of their conceptual schema is prerequisite for further learning. As they begin to understand other theories of reading, it is assumed that cognitive dissonance will motivate them to restructure their knowledge base. At this point it would be appropriate to set up demonstration situations to explore with the interns practices that are consistent with an interactive theory of teaching and reading. Frequent feedback to inform teacher interns about their consistency or lack of consistency in relating current theory to practice should be an integral component of the training model. Finally this process of examining theory and applying theory to practice needs to be extended over the entire program to optimize teacher interns' learning and long term retention of the training program content. Our aspiration for our teacher interns is that they will not only say what we say but will develop appropriate and effective teaching strategies.

REFERENCES

Adams, M. S., & Collins, A. *A schema-theoretic view of reading comprehension* (Tech. Rep. No. 32). Champaign, Illinois: Center for the Study of Reading, University of Illinois, April 1977. (ERIC Document Reproduction Service No. ED 144 039)

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Goodman, K. S. Reading: A psycholinguistic guessing game. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading.* Newark, Delaware: International Reading Association, 1976.

Goodman, Y. M. & Burke, C. L. *Reading miscue inventory.* New York: Macmillan Publishing Company, 1971.

Harste, J. C. & Burke, C. L. A new hypothesis for reading teacher research: Both the teaching and learning of reading are theoretically based. In P. D. Pearson (Ed.), *Reading: Theory, research, and practice.* Twenty-sixth Yearbook of the National Reading Conference. Clemson, South Carolina: The National Reading Conference, 1977.

Meyers, R. S. *Information processing as a tool for selecting instructional strategies.* Paper presented at The National Reading Conference, New Orleans, Louisiana, December 1978. (ERIC Document Reproduction Service No. ED 162 258)

Mitchell, K. A. *Patterns of teacher-student responses to oral reading errors as related to teachers' previous training in different theoretical frameworks.* Unpublished doctoral dissertation, New York University, 1978.

Ringler, L. H., Smith-Burke, M. T., & Meyers, R. S. *Focusing on reading models for teacher education.* Paper presented at The National Reading Conference, New Orleans, Louisiana, December 1977.

Ringler, L. H., Smith-Burke, M. T. & Meyers, R. S. *Reading comprehension.* Symposium presented at Seventh World Congress in Reading, International Reading Association, Hamburg, Germany, August 1978.

Ringler, L. H., Smith-Burke, M. T., & Meyers, R. S. Interactive model of teacher education. In M. Kamil & A. J. Moe (Eds.), *Reading research: Studies and applications.* Twenty-eighth Yearbook of the National Reading Conference. Clemson, South Carolina: National Reading Conference, 1979.

Rumelhart, D. E. *Toward an interactive model of reading* (Tech. Rep. No. 56). San Diego, California: Center for Human Information Processing, University of California, 1976.

Smith, F. *Understanding reading* (2nd ed.). New York: Holt, Rinehart, & Winston, 1978.

JAMES A. DINNAN
University of Georgia

AN EVALUATION OF LITERACY PROGRAMS FOR MATURE ADULTS

As a significant proportion of the population of the United States grows older, the ranks of the illiterate grow larger. "Illiterate" has many definitions but for this discussion it will be "the inability to function with reading material beyond the fifth grade level." The fifth grade level has been chosen because of the individual's need to integrate and associate the new data with past experiences and perceive the relationships presented by the author. Below this level the task is one of learning the operations of the reading code and the many pieces that comprise the whole, namely the act of receiving and processing whole bodies of knowledge in written form.

The mature adult shall be one who has attained the chronological age of 55 years or older. According to the National Center for Education Statistics (1972, 1975) this group has consistently comprised between 8 and 10 percent of the population of adult basic education programs. The percentage is somehwat smaller in private agencies.

The individuals who come to A.B.E. programs or continuing education programs have been identified by many agencies. "Who fits what programs?" depends upon the agency and its criteria for inclusion, and estimates range up to tens of million of people who supposedly need help for the "good" life. One factor that keeps eluding the researcher is the importance of human communication between the A.B.E. student and the teacher. Books, machines, movies, objects, even money, have been used in investigations, but what a teacher does with them, or for that matter without them, is the essence of that act which will open the world of the individual, it is not material contact, but *human* contact which will make the difference.

In order to assess the background of a group of A.B.E. teachers, to see what experiences they brought to the classroom, a survey was completed in an adult workshop involving 89 teachers. Basically, the results showed that three-fourths of the group were between 25 and 55 years old. The average teaching experience, mainly in the elementary schools, was over 10 years. Undergraduate preparation was generally in teacher education with an emphasis on the language arts and the social sciences. Results of this survey indicate that the adult basic education teacher is probably a woman about 40 years old who has over 10 years of regular teaching experience, usually in the elementary grades, who has been involved in adult education programs less than 4 years. The work in adult education is a secondary income source.

How our hypothetical teacher is teaching the adult groups, especially how she is teaching them to read, is not as clearly defined. Attendance itself at this workshop, even though rewarded with college credit and a stipend, may well indicate a less than satisfied attitude on the part of these teachers about their teaching skills. This group was well qualified for teaching, based on college degrees earned and classroom experience. The real question is what portion of the knowledge and experience of members of this pool could be directly applied to the special needs of the adult learner? A careful reading of the responses, indicates a lack of preparation for the teaching of adults. In some instances answers indicated a traditional elementary school curriculum was being inappropriately followed, however, it may well have been lack of knowhow as opposed to lack of motivation for change.

Because reading is such a large part of adult basic education, courses should be made available to these teachers that are designed to coordinate the expertise in reading methodology they already possess with the special charcteristics of adult learners; such as the use of their vast experience, extremely high motivation, job orientation, or physical limitations.

Specific programs for the illiterate mature adult are few and far between. The programs reported in the special issue on "Reading" in *Education Gerontology* (1979) all involve people who have mastered the fundamentals of reading. The public school programs such as the "adult college boom" after 65 is much in the news, but the programs for literacy are still very sparse. There are many situations where the mature adult is part of a group and the instructions become secondary to the social interaction that takes place before, during and after the programs. The goals of some programs are not in line with many of the ideas brought to the class by the mature adult nor are most of the materials suited to adults.

In summary, the state of literacy programs for mature adults is one of extreme neglect. A search of the last five yearbooks of the *National Reading Conference* (24th through 28th) revealed only two short articles on mature reading. Teacher preparation has been slow in preparation of courses for understanding and teaching adults in general. Methods and materials are often elementary school oriented and devoid of an awareness of adult performance needs. Research has been neglected in this area for years.

The future of the problem to be faced, with the reported performance of today's school population, is certainly grave unless more people are willing to put time and effort into understanding and solving the problems of Adult Learning.

REFERENCES

Adult Education (Adult Basic and Secondary Level Program Statistics). Washington, D.C.: National Center for Education Statistics, U.S. Department of H.E.W., 1972, 1975.

Educational Gerontology, 1979, *4*, (3).

Kamil, M. L. & Moe, A. J. (Eds.). *Reading research: Studies and applications.* Twenty-eighth Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1979.

McNinch, G. H. & Miller, W. D. (Eds.). *Reading: Convention and inquiry.* Twenty-fourth Yearbook of the national Reading Conference. Clemson, S.C.: The National Reading Conference, 1975.

Miller, W. D. & McNinch, G. H. (Eds.). *Reflections and investigations on reading.* Twenty-fifth Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1976.

Pearson, P.D. (Ed.). *Reading: Theory, research and practice.* Twenty-sixth Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1977.

Pearson, P. D. & Hansen, J. (Eds.). *Reading: Disciplined inquiry in process and practice.* Twenty-seventh Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1978.

MOLLY WILSON DROTTER
Idaho State University

PHYSICAL AND PSYCHOLOGICAL DECREMENTS AFFECTING READING AMONG THE AGED

Previous work has suggested that reading can be an important activity in the lives of the elderly. Reading can serve as both a leisure time activity and a means of obtaining information necessary for everyday living. It can provide a way to spend time alone or be the basis for a group experience (Lovelace, 1979; Wilson, 1979). Yet, despite all the value that reading can provide, few elderly people read (Kingston, 1979). While motivation is a contributor to this lack of interest, physical and psychological decrements associated with aging also affect reading behavior and partially explain the lack of interest in reading among the elderly. Researchers know that some of the decrements that are a part of the natural process of aging affect information processing and retrieval systems used in reading. The purpose in this paper is to discuss selected aspects of the important visual, perceptual, and cognitive changes that occur with age and affect the ability to read. Means for overcoming the decrements will be discussed, also.

DECREMENTS WITH AGE

Visual Changes

As part of the natural aging process, changes occur within the structure of the eye itself. The iris and pupil decrease in size with age; the lens becomes more opaque; the vitreous humor becomes thicker. Basically, three results of the physical changes in the eye affect the ability to read: 1) presbyopia (the inability to focus on near point); 2) higher threshold for illumination; 3) prolonged time needed for dark-light adaptation (McFarland, 1967). Less light is actually taken in by the aging eye and what is taken in is more diffused by the opacity and thickness of the eyeball itself. Thus, the elderly person finds it more difficult to focus on a page of print (Weale, 1965). The thickening of the lens to form cataracts has long been recognized as a visual problem in some elderly that requires special attention and compensation for reading. But all elderly people with or without cataracts need increased illumination and glare reduction when working with close materials because of the aging of the eye.

Perceptual Changes

In addition to physical changes within the eye, perceptual changes occur also. It takes longer with increased age for an image to register in the brain as having been seen (Weale, 1965). This is partly caused by the reduced intake of sensory data, which is the fault of the aging eye (McFarland, 1967); but, in addition, research indicates that neural transmission to the brain centers slows with age (Griew, 1963). Two other perceptual processes contribute to decrements with age also: 1) longer preparatory interval; 2) increase of both after effects and incidental neural activity (Welford, 1965).

A great deal of research has been done on the slowing of response time in the elderly (Botwinick, 1978). Current researchers agree that the slowing of response is due to a lengthening of preparatory interval in the central nervous system. This means that the elderly are slower to actually form a neural set for response, slower to attend (Rabbitt, 1965). So, response is slowed by the longer time taken by the elderly to integrate stimuli rather than in the actual psychomotor response movement (McFarland, 1967).

In addition to the lengthening of preparatory interval, both the amount of neural noise and the length of after effects increase with age (Welford, 1965). As the neurological perceptual system begins to lose the ability to discriminate necessary stimuli from noise, there is a loss of ability to attend to appropriate stimuli (Rabbitt, 1965). This contributes to a general perceptual confusion, which is now commonly explained by the Stimulus Persistence Theory. This theory states that in addition to increased general noise on the neural pathways, the elderly person's perceptual images persist longer in the central nervous system. Thus, in a task requiring response to a series of fast-paced images, the old after image interferes with the new image, increases confusion, and lengthens response time (Axelrod, 1963; Welford, 1965; Botwinick, 1978).

In summary, four perceptual decrements occur with age: 1) There is a decrease in amount of sensory data taken in by the aging eye. 2) An increase occurs in both the time to transmit a neural signal to the central nervous system and in integration time once it gets there. 3) Amount of neural noise and persistence of after image effects add to perceptual confusion. 4) Response time is increased for all tasks. In turn, these perceptual factors contribute to changes in cognitive processing abilities in the elderly.

Cognitive Changes

The cognitive changes associated with age have been the most widely researched areas in gerontology. At present, the consensus among researchers is that while short term memory ability decreases with age, there are few changes in long term memory capacity. In addition, it is suggested that the elderly lose some ability to do abstract problem solving tasks but retain both concrete problem solving abilities and verbal intelligence capabilities (Botwinick, 1978). An intelligence test given to adults aged 40 to 70 showed no decrease in information and comprehension subtests and an increase with age on vocabulary and similarities subtests (Bromley, 1963). This suggests that the basic intellectual abilities necessary for reading remain intact with age. However, two factors that may have an adverse effect on reading comprehension are 1) inability to integrate material and form a set for comprehension, and 2) decrease in ability to adapt, i.e. increased rigidity.

Experimental verbal learning tasks have shown that on a cognitive level, elderly fail to integrate information gathered from materials, to see patterns of similarities and differences in a problem, or to form a set for response. They find it more difficult to select the salient features or information in a task and to discard the irrelevant materials (Rabbitt, 1965; Arenberg, 1973). Thus, they are poorer than younger subjects at organizing a learning task independently (Botwinick, 1978).

In addition, people become more rigid with age thus finding it harder to change set once it is established. They will impose set on non-structured material (Riegel, 1965) and lump patterns or events into one category, thereby missing important differences and details (Welford, 1963).

Elderly people take longer to respond and to learn a task, but retain information as well as young people once they have learned (Arenberg, 1973). They can pick up speed on overlearned material (Riegel, 1965) but generally do best in self-paced situations (Welford, 1963) because of perceptual slowing and the cognitive factors mentioned here.

OVERCOMING THE DECREMENTS

Visual Prostheses

Great strides have been made mainly through the work of industry and medical science to compensate for the visual problems of the aged. New technology in the production of eyeglasses corrects presbyopia and increases acuity. Devices for increasing illumination are available, as are magnification devices. Advances in laser surgery make it easier and safer to remove cataracts at earlier stages. In addition, innovative methods of presenting written material are widely available. The Talking Books Program from the Library of Congress provides free materials to every library and private citizen in the United States. Books in large print and Braille books make it possible for the elderly, or any visually handicapped person, to continue reading despite the handicap. Therefore the natural aging of the eyes need not cause a decrease in reading.

Motivation and Interest

Motivation to read is the great means of compensation. While research shows that informaton intake, processing, and retrieval systems slow with age, the role of motivation in overcoming these and any physical decrements must not be minimized. Pleasure reading is, for the elderly, a self-paced task. Therefore, the slowing due to perceptual and cognitive changes is not of major consequence. Though being a slower reader than in earlier life may frustrate some, it is a personal frustration and not a factor causing inability to read. Motivation to read and the interest in materials presented can overcome problems caused by visual, perceptual, and cognitive decrements of aging.

This points to the fact that educators and parents need to train lifelong reading habits beginning in early childhood. For those who come to old age already loving to read, motivation can overcome the decrements of aging that could be hindrances. For those who are not already avid readers, the perceptual and cognitive slowing may make reading too laborious a task to be enjoyable. Experience and motivation can fully compensate for the decreased biological capacity associated with age (Griew, 1963), not only in reading but in all areas.

In addition, the key roles of experience and motivation as compensators points out the need for lifelong educational programs. Continuing education can keep people motivated and contribute to continued expansion of the experience base that is necessary to overcome biological degeneration (Baltes & Labouvie, 1973). Research into the types of materials that interest the elderly and the best formats for presentation of reading materials is needed.

Need for Further Research

The gerontological research on perceptual and cognitive decrements is abundant. Yet until very recently all of this research had been in an experimental, laboratory setting. Laboratory studies are unnatural tasks, and arouse a great deal of anxiety in the elderly (Riegel, 1965; Botwinick, 1978). It is impossible to separate learning from performance in any study. With the elderly anxiety about working in a laboratory setting on an unusual verbal task may affect performance to such an extent that it is not an accurate measure of ability to learn.

In addition, Botwinick (1978) states that in the available gerontological research on which all the conclusions about decrements occurring with age are based, educational level has never been sufficiently controlled. Many of the effects found with the elderly may be a function of the education level of the elderly samples used in research comparing young and old samples rather than true decrements of age.

Reading may present a way to further study verbal learning in the aged. Experimental psychologists (e.g., Rabbitt, 1977) acknowledge the fact that reading studies could make a contribution to knowledge about verbal learning behavior of the elderly.

Reading is a thought process, and as such, calls on a number of verbal learning skills, not just the isolated variables tapped in experimental verbal learning studies. The task of reading combines what is on the page with the reader's background experience and thought set. It is a more natural task than paired associates or serial learning tasks. Existing reading tests can provide more information on educational ability level, also.

While experimental studies can shed some light on the verbal behavior of the elderly, reading is a relatively untapped area of verbal learning that can add to the complete profile of elderly verbal behavior (Arenberg, 1973). In addition, researchers have seemingly overlooked the contribution that using an elderly sample may make on the study of reading comprehension. The elderly comprise a special group in society. The use of them in prose learning studies could shed light into the effects of motivation, adaptability, and textual organization on overcoming physical and psychological factors that should affect reading comprehension.

CONCLUSIONS

While reading has been recognized as a potentially useful and enjoyable pastime for the elderly, physical and psychological decrements affect the ability of the elderly to read. The eyes age, and near-point tasks become more difficult. In addition to reduced sensory intake, perceptual changes occur. The central nervous system slows; thus, data travels more slowly to the brain and is integrated more slowly once it arrives. Persistence of after images and increased noise in the neural system add to confusion and slow reaction time. While neither verbal abilities nor long term memory show much decline in the elderly, difficulty in forming a thought set, and rigidity in using ineffective strategies affect cognitive processing abilities.

Several means of overcoming visual decrements are available. Eyeglasses, advanced methods of surgery, and specially designed reading materials are all available to compensate for the aging of the eyes. The factors of motivation and experience are crucial in overcoming perceptual and cognitive deficits.

Finally, further research involving the elderly in reading studies is necessary. A number of the decrements found in experimental verbal learning studies may not affect performance on the more natural task of reading due to interest in and meaningfulness of material. In addition, the use of elderly subjects in studies of reading can further knowledge of the effects of motivation, experience, and textual organization on comprehension.

REFERENCES

Arenberg, D. Cognition, and aging: Verbal learning, memory and problem solving. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging.* Washington, D.C.: American Psychological Association, 1973.

Axelrod, S. Cognitive tasks in several modalities. In R. H. Williams, C. Tibbitts, & W. Donahue (Eds.), *Processes of aging* (Vol. 1). New York: Atherton Press, 1963.

Baltes, P. B., & Labouvie, G. V. Adult development of intellectual performance: Description, explanation, and modification. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging.* Washington, D.C.: American Psychological Association, 1973.

Botwinick, J. *Aging and behavior.* New York: Springer Publishing Co., 1978.

Bromley, D. B. Age differences in conceptual abilities. In R. H. Williams, C. Tibbitts, & W. Donahue (Eds.), *Processes of aging* (Vol. 1). New York: Atherton Press, 1963.

Griew, S. Information transmission and age. In R. H. Williams, C. Tibbitts, & W. Donahue (Eds.), *Processes of aging* (Vol. 1). New York: Atherton Press, 1963.

Kingston, A. J. Reading and the aged: A statement of the problem. *Journal of Educational Gerontology,* 1979, *4,* 205-208.

Lovelace, T. L. Reading activities to enhance the lives of nursing home patients. *Journal of Educational Gerontology,* 1979, *4,* 239-244.

McFarland, R. A. The sensory and perceptual processes in aging. In K. W. Schaie (Ed.), *Theory and methods of research on aging.* Morgantown, W. Va.: West Virginia Press, 1967.

Rabbitt, P. M. A. Age and discrimination between complex stimuli. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system.* Springfield, Ill.: Charles C. Thomas Publishers, 1965.

Rabbitt, P. M. W. Changes in problem solving ability in old age. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging.* New York: Van Nostrand Reinhold Co., 1977.

Riegel, K. F. Speed of verbal performance as a function of age and set. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system.* Springfield, Ill.: Charles C. Thomas Publishers, 1965.

Weale, R. A. On the eye. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system.* Springfield, Ill.: Charles C. Thomas Publishers, 1965.

Welford, A. T. Social, psychological and physiological gerontology: An experimental psychologist's approach. In R. H. Williams, C. Tibbitts, & W. Donahue (Eds.), *Processes of aging* (Vol. 1). New York: Atherton Press, 1963.

Welford, A. T. Performance: Biological mechanisms of age: A theoretical sketch. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system.* Springfield, Ill.: Charles C. Thomas Publishers, 1965.

Wilson, M. M. Enhancing the lives of the aged in a retirement center through a program of reading. *Journal of Educational Gerontology,* 1979, *4,* 245-252.

KATHLEEN M. NGANDU
Hood College

BILL O'ROURKE
Colorado Springs Public Schools

READING ATTITUDES, HABITS, INTERESTS, AND MOTIVATIONS OF THE ELDERLY

Research in reading typically focuses on children, young and middle-aged adults. We have prolific information concerning these populations. In contrast, very little attention is paid to the older-aged groups. A recent computer research of *Dissertation Abstracts International* as well as *ERIC* sources located less than a dozen data based studies focusing on reading attitudes, habits, interests, and/or motivations of the elderly.

De Santi's (1976) research identified some of the reading strategies, interests, and habits of four older adults. Wolf (1975) specifically examined leisure-time reading behaviors. Mitzer's (1975) research examined reading differences between active and educationally non-active older persons. And an earlier study by Hoar (1960) looked at general reading behaviors of the aged.

Romani (1973) and Lovelace (1977) briefly discussed reading interests of the elderly. Kingston (1977) outlined preferred reading topics of middle and upper-income older adults as well as their reading habits. Finally, a more comprehensive study by Moshey (1972) included types of reading materials a group of 50 retired adults generally selected.

Recognizing the need for more information about reading as related to older Americans, The University of Nebraska Center on Aging funded the seminal study discussed in this report. The specific purpose of this research was to answer the following questions as they relate to elderly individuals:

1. What are their attitudes about reading?
2. What types of reading materials are most popular?
3. What are their most frequent sources of reading materials?
4. How much time is spent in daily reading, as compared with television viewing?
5. What are their preferred reading interests?
6. What motivates individuals to read (or not read)?

METHOD

This case study dealt with persons attending the Senior Diners program in a medium-sized Nebraska city. In order to encourage social contact, the program serves a nutritious meal to older persons in local community centers. These persons live in independent, rather than retirement home situations. Most persons are bussed in Senior Diner vans, although some drive cars or walk to the centers.

The researchers and their assistants visited all nine Senior Diner sites where they requested volunteers to complete a 10 item, one-page questionnaire. Two hundred sixty-seven persons, about 80% of the total group, completed the form. The respondents' ages ranged from 60-93 years, averaging 71 years. Sixty-seven percent were female and thirty-three percent were males.

Persons answered the questionnaire in small groups or individually with a research assistant if they needed help to write their answers or seemed to desire extra personal attention. Most of the subjects took about ten minutes to answer the questions. Generally the respondents were cooperative and several persons talked extensively about information they had written on their forms.

The questionnaire included a total of eight closed and open-ended questions that focused on the original research questions. Answers were later tallied and are summarized in the following section.

RESULTS AND DISCUSSION

Attitudes

A positive attitude about reading was quite evident. Given five attitude choices, "love reading" was selected by 32% of the group, "like reading" was chosen by 55%, "no feelings" was indicated by 10%, only three percent said they "disliked reading," and no one checked "hate reading."

Elderly persons enjoy reading and this feeling should be enhanced whenever possible, as reading activities are productive ways to fill leisure-time hours.

Popular Reading Materials

Elderly adults most frequently read newspapers (76% of all subjects indicated they often read this material) and magazines (63%). To a lesser extent they read books (38%) and pamphlets (20%). The *Bible* was specifically identified as the main reading material by 19% of the group and was therefore not included in the general category of books. Other items were mentioned by only one percent or less of the group.

Perhaps newspapers and magazines are generally preferred due to their short, timely articles as well as the ease of obtaining them through subscription. Currently there are many special interest publications for a variety of special interest groups, such as working women, sports fans, and singles, but there are very few magazines specifically designed for the elderly. This gap must be filled.

Sources of Reading Materials

The major reading sources for the Senior Diners were subscriptions (59% of all persons said they subscribed to at least one publication), churches (29%), friends (27%) and senior centers (23%). Libraries (19%), family (14%) and bookstores (7%) were rated lower, while bookmobiles (2%) and newsstands (1%) were checked by very few people. Other materials such as personal letters and advertisements provided nine percent of the sources.

The best sources of materials seem to be those that come via home delivery, churches, or senior centers. Since these materials are most readily available, they are logical avenues. It seems that there could be an increased use of public libraries; however, transportation would have to be arranged in many cases.

264

Time Spent Reading and Viewing TV

The total group averaged 2.4 hours of daily reading with a range from "no time" to six hours. Persons were also asked to note the amount of time they spent watching television. Interestingly, the average time spent viewing TV was also 2.4 hours for the total group, but the range extended from "no time" to nine hours.

Reading is an important daily activity for many older persons. Compared to adults in general, they read slightly less than the average 2.7 hours of daily job-related and free-time reading (Mikulecky, 1978). It must also be emphasized that there are several elderly who read much less than this, possibly because of difficulties with eyesight.

Although several persons complained about the poor quality of television programs, this medium still competes for their attention. According to the results of a two-year project at the Annenberg School of Commerce, University of Pennsylvania, "Older people are . . . badly shortchanged in television drama"—they are portrayed very negatively (Gerbner, as cited in Kernan, 1979). Perhaps the immediate accessibility of the TV, compared with the greater difficulty in obtaining reading materials, influences some older Americans' decisions to watch television.

Preferred Reading Interests

When asked to identify their favorite reading interests in an open-ended question, the Senior Diners gave a priority to religion (43% of all persons supplied this topic) and news (18%). The topic of history was listed by five percent of the group, while the sports topic was provided by four percent. Biography and travel were mentioned by three percent of the cases, while politics, crafts, medicine and home-making topics were selected by two percent in each case. Finally, one percent or less mentioned the topics of animals/nature/plants, comics, finance, poetry, science, education, "good old days," and psychology as their preferred reading topics.

There is a great diversity among preferred reading topics, but religion and the *Bible* were continually emphasized by many persons. It therefore seems logical that church groups promote reading clubs, book sharing sessions, or book lending libraries to encourage reading pursuits of the elderly.

Motivation for Reading (or Not Reading)

These persons said they read mainly because they enjoy it (23% of all subjects stated this reason) and wanted to be informed or educated (23%). Several others read mainly to pass the time (20%). Two percent read only for religious reasons, while one percent or less read because they were lonesome, to relax, to check for coupons, or because they couldn't watch television.

Twenty-six percent of the Senior Diners stated that there were times when they chose not to read because of eye difficulties. Two percent felt they did not have time to read, lacked interest, or wanted to do other things. One percent or less wrote that they did not care to read, were too sleepy, had ill health, or lacked reading abilities.

Difficulties of old age, mainly eyesight, have a negative influence on persons' reading habits. In cases where individuals can no longer read well, large type or recorded materials may help. Talking books are provided free of charge by organizations such as The Jewish Guild for the Blind (Cassette

Library, The Jewish Guild for the Blind, 15 W. 65 Street, New York, N. Y. 10023). Older persons could also be read to by others. For example, in this particular city, a local scout troop read to individuals in a nursing home. This same type of service is needed by those in independent-living situations.

It may also be the case that elderly persons do not have sufficient home lighting to read well, as a recent medical study (Davis, 1979) concluded. Fifty percent of the older persons attending a sight disability clinic could read much better in the clinic than they could at home. Upon closer investigation it was found that their home lighting was typically only ten percent of that at the clinic. The researchers suggested that simply adding a 60 watt bulb and a six dollar lamp would greatly improve the elderly's vision at home.

CONCLUSIONS

The older American represents a growing minority whose concerns must be recognized, especially by community organizations, recreational planners, and other groups that service them. Yes, financial outlays, cooperative planning, and a genuine interest will be needed to support some of the previously made suggestions; however, reading activities for elderly persons must be encouraged. We believe that individuals who read are persons who think, remain mentally active, and generally feel satisfied with this activity. These benefits will certainly directly or indirectly benefit our total society.

The efforts of this initial seminal study need to be expanded to larger as well as more diverse populations. For example, rural community elderly, or those living in retirement homes could be examined. Also, what common factors have influenced their reading? If the backgrounds of avid life-long readers could be studied, a "reading autobiography" might help isolate particular variables which seem consistent among these older persons.

As we add information about elderly Americans' reading attitudes, habits, interests, and motivation to that which we already have about children and adults, we should be able to understand more completely the developmental reading process. Thus, the span from the very young to the very old will be completed.

REFERENCES

Davis, S. *Speaking on health.* Commentary on radio station WTOP, Washington, D. C.: November 15, 1979.

De Santi, R. J. The older reader: An investigation of the reading strategies, habits and interests of four persons sixty years of age or older (Doctoral dissertation, Indiana University, 1976). *Dissertation Abstracts International,* 1976, *37* (08-A), 4719. (University Microfilms No. 77-3285)

Hoar, J. R. Reading, listening, and viewing behaviors of the aged: An inventory of the mass communication habits and preferences of 200 aged persons in Oxford, Mississippi (Doctoral dissertation, University of Iowa, 1960) *Dissertation Abstracts International,* 1969, *20* (12), 4649. (University Microfilms No. 60-01556)

Kernan, M. Old age: TV tells it wrong. *The Washington Post,* September 18, 1979, p. C3.

Kingston, J. J., Jr. *Reading and the aged: A statement of the problems.* Paper presented at the Annual Meeting of the International Reading Association, Miami, May 1977.(ERIC Document Reproduction Service No. ED 137 739)

Lovelace, T. *Enhancing the lives of nursing home patients through reading activities.* Paper presented at the Annual Meeting of the International Reading Association, Miami, May 1977. (ERIC Document Reproduction Service No. ED 140 274)

Mikulecky, L. J. Overview of a survey of adult reading habits, attitudes and motivations. *ADSIG Journal,* 1978, *1,*(2), 24-31.

Mizer, M. M. Differences between educationally active and noneducationally active older adults in reading (Doctoral dissertation, Arizona State University, 1975). *Dissertation Abstracts International,* 1975, *36* (05-A), 2507. (University Microfilms No. 75-24225)

J. MICHAEL BENNETT
DAVID M. WARK
University of Minnesota

TEST ANXIETY OF COLLEGE STUDENTS — TEST OF A MODEL

Test anxiety is as old as testing, but serious attempts at definition and explication have been in development for only twenty years or so. The results of strong test anxiety are observed in several different, though related, ways. There are some persons who become physically ill when confronted with a test—particularly one, the results of which, if low, may promise to bar the door to some especially desirable circumstance. A man of the authors' acquaintance would frequently forget his age and often misspell his name when placed in the testing situation.

Another manifestation of the severe anxiety spectrum is the student who always performs better in every school situation than he or she does on a formalized test. Similarly, the person who suddenly remembers all the material covered by a test, after the test has been handed in, is probably suffering from acute test anxiety, as are those persons who experience unaccountable gaps in comprehension in the middle of an essay, or even a paragraph. The highly anxious person, therefore, does not experience the excitement and rush of adrenalin which helps one to perform at the highest levels. Rather, test anxiety identifies a set of phenomena which debilitate and incapacitate.

RATIONALE

The relationship between test anxiety and college achievement has been reported by several researchers. Gaudry and Spielberger (1971) reported that in general, highly anxious students at each level of ability were far more likely to drop out of college than were low anxious students with the same level of intellectual competence. Sarason (1972) demonstrated that highly anxious students do as well as low anxious peers when placed in non-threatening or non-test-like situations. The problem, of course, is that what happens in a college classroom is frequently competitive and test-like for reasons of economy, efficiency, and tradition. Therefore, the highly anxious student may be sharply disadvantaged.

Of the several studies on reading and anxiety (Holmes, 1972; Johnson & Hummel, 1971; Kestenbaum & Weiner, 1970; Lombardi, 1974; Proger, 1973) none focus on the university level student. All report studies of elementary or secondary readers. The extant literature shows a clear negative correlation between anxiety and various measures of comprehension.

Recent efforts have been directed toward developing and validating an explanation of the impact on test anxiety on cognitive activity. Muller (1976) employs a cue utilization hypothesis which may have implications for students at all ages. Muller holds that anxious or aroused students may miss some clues in coding stimuli for memory. The calm student learning a list of words notices, in addition to meaning, such characteristics as shape, initial and final letters, rhymes, and emotional loadings. For the anxious student, some of these labels for sorting and retrieving are not effective. The consequence is that information or words read may not be available for recall. In addition, highly anxious persons seem to be largely unaware of associations or "clustering" of related words contained in the learning task. In a reading situation the learner must cluster information obtained in various parts of the paragraph in order to make sense of the entire paragraph.

It seems probable, therefore, that if Muller's findings are applicable, a highly anxious college student would perform satisfactorily on receptive or "detail" test items, and less well on reflective or "thought" questions. Receptive questions require one to have noticed a certain bit of information, and the "tunnel vision" of the highly anxious should not detract overmuch from acquisition of these facts. Reflective questions, on the other hand, assume the association of ideas or events reported in several places in a passage, and may even require inferences or judgments based on such association or clustering of both explicit and implicit information.

The investigators decided to attempt a comparison of High Anxiety and Low Anxiety university students in an effort to 1) extend the available literature on reading and test anxiety to adults, 2) compare two well-known college level reading tests in terms of their sensitivity to anxiety effects, and 3) examine Muller's cue utilization model applied to receptive versus reflective comprehension at the university level.

POPULATION

The subject pool for the study was the group of students enrolled in the fall quarter sections of Efficient Reading at a large public university. Of the 100 plus students, eighty-four (84) completed all testing. Approximately fifty per cent were female. Slightly over half the subjects were freshmen, with the remaining ranging from sophomores to graduate students.

PROCEDURES AND RESULTS

All subjects were given the *Test Anxiety Scale* (Sarason, 1958). The observed range of response was from zero (no anxiety) to 14 of 16 possible points. The mean was 6.630, with a standard deviation of 3.247, giving the distribution a slight positive skew. The 27 students who ranged from zero to four on the scale were classed as Low Anxiety. A High Anxiety rating was reserved for the 25 subjects whose scores ranged from nine to 14. The remaining 32 students were classified as Medium Anxiety. It should be noted that students who enroll in these courses are able students with above average reading ability and are, therefore, not especially nervous or anxious as a group.

The subjects were then given two reading tests to find out if there were different levels of comprehension associated with the three classifications of test anxiety. On the *Nelson-Denny, Form C,* comprehension section, the high anxious students scored at the 23rd percentile, the medium anxious at the 47th percentile, and the low anxious at the 50th. The ANOVA yielded $F(2,81)=6.10, p<.01$. The differential results were similar for those taking the

McGraw-Hill, Form A. On the comprehension section the high anxious scored at the 50th percentile, the medium anxious scored at the 73rd percentile, and the low anxious subjects averaged the 87th percentile. For these data, the ANOVA showed $F(2,81)=9.05$, $p<.01$. For both tests, the largest separation was between the medium and high anxious categories. It seems obvious that test anxiety is a potent factor in student reading performance.

It seems worth mentioning here that the investigators were surprised at the results, in that previous research (Wark, Bennett, Wolinski, 1980; Bennett, Wark, 1980) indicated that the *Nelson-Denny* was insensitive to test anxiety. It was our previous opinion that this was true because of the progressive difficulty and, especially, the "speededness" of the *Nelson-Denny*—factors which act as "leveling agents" and tend to iron out anxiety effects. However, prior research compared two groups of students who took one or the other test. The current study is the first in which students took both instruments.

The data can be used to test the cue use model. Muller's work would suggest that differences between the comprehension of high and low anxiety students would be a function of item types. For detail questions (which are answered on the basis of a particular fact found in a particular place in the text) there should be relatively little difference in the performance of high and low anxiety students. On the other hand, for reflective questions (that require inference and judgment), there may be differences. The high anxious students would have difficulty keeping relevant facts in mind, clustering them, seeing their association, and then coming to an independent conclusion. These reflective questions, which are answered based on the material from various places in the reading paragraph, might discriminate between high and low anxious students. If the expected differences appeared for reflective questions, but not for detail questions, the results would strengthen the cue utilization model of test anxiety.

Items from the *McGraw-Hill* and *Nelson-Denny* tests were categorized, where possible, into reflective or detail groups. For the *Nelson-Denny,* classification of items was as specified in the test manual. The *McGraw-Hill* items were categorized by a three-judge panel. (See Wark, Bennett & Wolinski, 1980.) Not all questions could be categorized.

There are certain findings which support the test of the cue utilization model. First, high anxiety students tend to score lower (M=.682) on both types of items than do students (M=.783) who are less anxious. Second, reflective questions generally are more difficult (M=.665) than detail questions (M=.800). We would expect the first finding on the basis of the more general relationship between anxiety and reading comprehension. The critical test for the cue use hypothesis, however, is the difference between the high and low anxious students for detail questions versus the same difference for reflective questions. For both tests the over-all difference is greater for reflective items (M=.133) than for detail items (M=.07), as predicted by the hypothesis. Any value of the ratio of reflected/detailed comprehension greater than 1.0 would be supportive. For the *McGraw-Hill* the test ratio is 2.37; for the *Nelson-Denny* it is 1.63. Apparently the *Nelson-Denny* test is less sensitive to anxiety.

While the general direction of all differences is that predicted by the model, the size of the difference does not reach conventional statistical significance levels. The pooled test for detailed versus reflective items, across both tests and both levels of anxiety is $t=.922$, df=76, $p<.17$. The cue utilization model gets some, although weak, support from these data.

These findings can be regarded as tentative. If they are replicable, they will have relevance for developing strategies for compensatory or corrective teaching and counseling methodology. These strategies should grow out of an increased understanding of what test anxiety is, and how it is caused and manifested.

REFERENCES

Bennett, M., & Wark, D. Some effects of anxiety on university students. In J. Fisher (Ed.), *North Central Reading Association Yearbook* (Vol. 10), 1980.

Brown, J. I. *Examiner's Manual — The Nelson-Denny Reading Test, Forms C and D.* Boston: Houghton Mifflin Company, 1973.

Brown, J. I., Bennett, J. M., & Hanna, G. S. *The Nelson-Denny Reading Test for High School and College, Forms E and F.* Iowa City: The Riverside Press, 1980

Gaudry, E., & Spielberger, C. D. *Anxiety and Educational Achievement.* Sidney: John Wiley & Sons, Australasia PTY, Ltd., 1971.

Holmes, C. C. *Specific effects of test anxiety on reading comprehension as measured by the cloze procedure.* Unpublished doctoral dissertation, University of Georgia, 1972.

Johnson, G. L., & Hummel, T. J. The effects of three models of test administration on the reading achievement scores of fifth graders. *Elementary and School Guidance and Counseling*, 1971, *6* (1), 21-26.

Kestenbaum, J., & Weiner, B. Achievement performance related to achievement motivation and test anxiety. *Journal of Consulting and Clinical Psychology*, 1970, *34* (3), 343-344.

Lombardi, V. A. *The effects of test anxiety on the relationship between intelligence and reading achievement in a group of seventh-grade suburban children.* Unpublished doctoral dissertation, Temple University, 1974.

McLaughlin, B. Intentional and incidental learning in human subjects: The role of instructions to learn and motivation. *Psychological Bulletin*, 1965, *63*, 359-376.

Muller, J. H. Anxiety and cue utilization in human learning and memory. In M. Zuckerman & C. C. Spielberger (Eds.), *Emotion and anxiety.* New York: Lawrence Erlbaum Associates, 1976.

Proger, B. B. Advance and concurrent organizers for detailed verbal passages used with elementary school pupils. *Journal of Educational Research*, 1973, *66* (10), 451-456.

Raygor, A. I. *Manual for the reading test: McGraw-Hill basic skills system.* Atlanta: McGraw-Hill Company, 1970.

Sarason, I. G. Inter-relationships among individual difference variables, behavior in psychotherapy, and verbal conditioning. *The Journal of Abnormal and Social Psychology*, 1958, *56*, 339-344.

Wark, D., Bennett, M., & Wolinski, J. The effects of anxiety on reading comprehension. In J. Fisher (Ed.), *North Central Reading Association Yearbook.* (Vol. 10), 1980.

Wildemuth, B. M. Test anxiety. An extensive bibliography. *ERIC/TM Report*, 1964. ERIC Clearing House on Tests, Measurement, and Evaluation. Princeton, New Jersey: Educational Testing Service, 1977.

Wine, J. Test anxiety and direction of attention. *Psychological Bulletin*, 1971, *76*, 92-104.

RONA F. FLIPPO
Georgia State University

THE NEED FOR COMPARISON STUDIES OF COLLEGE STUDENTS' READING GAINS IN DEVELOPMENTAL READING PROGRAMS USING GENERAL AND SPECIFIC LEVELS OF DIAGNOSIS

There is not sufficient research in the field of postsecondary reading to provide an answer to the question of how best to measure skills, diagnose needs, prescribe instruction, and remediate reading problems. A comprehensive review (Flippo, 1979) of the research from 1935 through 1978 reveals only two studies, Cox and Swalm (1973) and Smith (1974, 1977), that have compared the effects of prescriptive instruction based on specific diagnosis of diagnostic testing of reading abilities with effects based on general diagnosis or survey testing of reading abilities. The findings of Cox and Swalm significantly favored the results of instruction based on more specific diagnoses, while the findings of Smith indicated no significant difference between the results of instruction based on more specific diagnoses and the results of instruction based on general diagnoses of reading abilities.

In his paper which discusses the trends in college training, Raygor (1971) stated, "The cost-effectiveness factor is getting more and more important and we're going to have to start arguing for innovations in instructional programs on the basis of their ability to do a good job and to save money. College basic skills programs are going to be increasingly subject to the demand for cost-effectiveness considerations" (p. 7). Research investigating the most effective ways of testing, prescribing for and instructing postsecondary students in reading is necessary for a body of research that could be examined to meet cost-effectiveness considerations.

More particularly, research is needed to explore the comparison between diagnostic and survey testing in postsecondary reading. Although studies upon studies have been completed in postsecondary reading, (see Flippo, 1979, for review of research), we still do not know the answers to the seemingly simplest questions concerning the most effective and efficient methods of testing, diagnosing, and remediating the developmental or remedial postsecondary population we serve. Ironside (1969) called for research that would indicate whether diagnostic tests could generate more effective programs for college students than would survey tests. Yet Smith (1974), in her dissertation, states that she could not find any studies making that type of comparison. Although the Cox and Swalm (1973) and Smith (1974, 1977) studies are the beginnings of a body of research in the proposed area of study, they have contradictory findings and had limitations in the relatively small sizes of their total groups, n = 98 and n = 75, respectively.

The purpose for this study was to measure the effect of developmental reading instruction for postsecondary students prescribed from the specific results of a diagnostic reading test and to compare those results with the results of developmental reading instruction for postsecondary students prescribed from the general and more traditionally used results of a survey reading test. This study has increased the small body of already cited research on the comparison of general and specific diagnosis of postsecondary students' reading abilities.

Subjects

The sample in this study consisted of 226 students from an original group of approximately 450 freshmen enrolled in 15 different sections of "Effective Reading" in the fall semester 1978 at the University of South Carolina's College of General Studies, Columbia, S.C. These students were required to take the three credit course because of low Stanford Achievement Test scores.

Procedures

Students were randomly assigned to experimental (n = 112) or control (n = 114) treatments. The nine instructors have approximately an equal number of experimental and control students in each of their sections. Experimental and control treatments were provided by graduate assistants supervised by the researcher during lab times. All students wre treated with two hours per week of classroom instruction with their instructor and one hour per week of reading lab instruction. This lab instruction was based on each student's individual prescription. All students were pretested with two standardized testing instruments. On the basis of information from the subtests of the diagnostic instrument (*Stanford Diagnostic Reading Test*, Blue Level, 1976), individual prescriptions for reading skill improvement were written for the experimental students. Only the information from the subtests of the survey instrument (*Nelson-Denny Teading Test*, 1973) was used in writing prescriptions for the control students. The control students' Stanford Diagnostic answer sheets were left unscored until the end of the term. The original diagnoses from the pretest information on students in both treatments were not altered in any way throughout remediation. Both the control and experimental students had the same treatment, except for the amount of diagnostic information available for the initial counseling sessions to prescribe instruction and materials. Students in the experimental group were prescribed materials matched to their specific needs as measured by the eight subskills of the diagnostic test. Students in the control group were prescribed materials matched to the general skill areas as measured by the survey test. The rate subtest of the *Nelson-Denny* was used in prescriptions for control students, but not as a dependent variable in this study.

The prescription of the two groups was accomplished by using a card system designed by the researcher especially for testing instruments and materials selected for this study. The card system was programmed by the researcher, who cataloged all the materials into one or more of the 10 selected skill designations, and by difficulty level. (See Flippo, 1979, for more details.)

At the end of the semester, all students were posttested with different forms of the *Stanford Diagnostic Reading Test* and the *Nelson-Denny Reading Test*. The gain in the 10 selected subskills of reading was measured by the pretest and posttest differences of the two groups. Individual *t* tests, assuming equal variances, were used to test each of the 10 null hypotheses (one for each subskill, at $\propto = .05$) to determine whether there was a significant differences between the experimental and control groups on any of the subskills.

RESULTS

The means of the gain scores of the experimental students was significantly higher than the means of the gain scores of the control students in the subskills of literal comprehension, inferential comprehension, word meaning, phonetic analysis, structural analysis, and scanning and skimming. There was not a statistically significant difference between the means of the gain scores of the experimental students and the control students in the subskills of word parts, fast reading, vocabulary, and comprehension. However, on the subskills of fast reading and vocabulary, the observed gain of the experimental group was higher, but not substantially higher at the .05 level. (See summary table showing comparison of raw score gains on each of the subskills.)

This study's findings generally support the conclusions of Cox and Swalm (1973) that postsecondary students do better in developmental reading programs that utilize more specific diagnosis and prescription than in reading programs which use more general testing of reading abilities. However, in this study, as in the Smith study (1974, 1977), the gains made on the subskills from the *Nelson-Denny Reading Test* (vocabulary and comprehension) did not reflect evidence of a significant difference. Smith measured reading gains by the pretest and posttest differences of two groups on the *Nelson-Denny* alone. If this present study had solely used the *Nelson-Denny* to measure the gains, the findings would have supported the conclusions of Smith rather than that of Cox and Swalm. It appears that gains made on diagnostic reading tests may not necessarily be reflected in the results of survey reading tests.

This suggests that the *Nelson-Denny Reading Test* does not consistently appear to be a good discriminator of students that have improved more and students that have improved less. If Smith had given the diagnostic tests as pre and posttests to both her groups and compared those results between groups, she might have found that the group with more diagnostic information available before instruction did better as measured by the diagnostic tests. Unfortunately, this is not known, as Smith pretested both her groups with the *Nelson-Denny* and the more diagnostic tests, but posttested her groups with only the *Nelson-Denny*. Additionally, the gains on the subskill of comprehension, as measured by the *Nelson-Denny*, do not reflect the gains made on the subskills of literal and inferential comprehension, as measured by the *Stanford Diagnostic*. The *t* values of literal and inferential comprehension indicated that there was a significant difference between the gains of the two groups. Yet, there was no significant or observed difference between the means of the groups on comprehension as measured only by the *Nelson-Denny*.

IMPLICATIONS

The educational implications from the results of this study are that: (1) It is advisable to give postsecondary developmental reading students a diagnostic reading test to identify their reading strengths and weaknesses prior to prescription of their reading instruction. (2) Postsecondary students' awareness of their specific reading skill needs might better enhance chances of development in these skill areas. (3) Postsecondary reading program directors and instructors should be cautious of using a survey reading test instrument alone to evaluate the progress of their developmental students, as the survey instrument might not reflect growth and changes in the reading skills.

Given the high cost of developmental programs, the varied needs of the postsecondary students assigned to them, and the brief (usually one semester) time for remediation, it appears that a cost-effective program would be one that individualizes instruction as specifically as possible to individual students' reading skill needs. It is recommended that those responsible for the planning and implementation of these programs find ways to be as specific and individual in that instruction as possible and utilize the opportunity to conduct carefully controlled studies to contribute to research in this area.

Summary Table of *t* Tests Comparing Raw Score Gains

In all cases, n for experimental group = 112
and n for control group = 114

Subskills/Test	**Group**	**Mean Gain**	**Standard Deviation**	**t (* = $p < .05$)**
Literal Comprehension	Experimental	3.517	2.910	4.140*
(SDRT)	Control	1.578	4.030	
Inferential Comprehension	Experimental	2.857	2.667	5.526*
(SDRT)	Control	.614	3.385	
Word Meaning	Experimental	2.375	2.812	5.594*
(SDRT)	Control	-.096	3.753	
Word Parts	Experimental	2.357	3.221	.541
(SDRT)	Control	2.596	3.422	
Phonetic Analysis	Experimental	2.205	4.649	4.394*
(SDRT)	Control	.070	3.654	
Structural Analysis	Experimental	2.517	4.024	2.124*
(SDRT)	Control	1.447	3.540	
Scanning and Skimming	Experimental	3.473	3.193	6.401*
(SDRT)	Control	.061	4.668	
Fast Reading	Experimental	3.133	3.588	1.522
(SDRT)	Control	2.368	3.956	
Vocabulary	Experimental	4.526	3.815	1.623
(NDRT)	Control	3.675	4.062	
Comprehension	Experimental	5.883	4.699	.056
(NDRT)	Control	5.842	6.290	

REFERENCES

Brown, J. I., Nelson, M. J., & Denny, E. C. *Nelson-Denny Reading Test, Forms C and D.* Boston: Houghton Mifflin Company, 1973.

Cox, G. L., & Swalm, J. E. A formalized diagnostic approach to study skills instruction for specially-admitted college freshmen. In P. L. Nacke (Ed.), *Program and practices for college reading* (Vol. 2) Twenty-second Yearbook of the National Reading Conference. Boone, N.C.: National Reading Conference, 1973.

Flippo, R. F. *Comparison of college students' reading gains in a developmental reading program using general and specific levels of diagnosis.* Unpublished doctoral dissertation, University of Florida, 1979.

Ironside, R. A. *Who assesses reading status and progress-tests, teachers, or students?* Paper presented at the Twelfth Annual Meeting of the College Reading Association, Boston, Massachusetts, March 1969. (ERIC Document Reproduction Service No. ED 031 374)

Karlsen, B., Madden, R., & Gardner, E. F. *Stanford Diagnostic Reading Test, Blue Level, Form A and Form B.* New York: Harcourt Brace Jovanovich, 1976.

Raygor, A. L. Present and future trends in college reading and higher education: 1970 presidential address. In F. P. Greene (Ed.), *Reading: The right to participate.* Twentieth Yearbook of the National Reading Conference. Milwaukee: National Reading Conference, 1971.

Smith, B. D. The effect of differential diagnosis upon reading improvement for high risk junior college students (Doctoral dissertation, Georgia State University, 1974). *Dissertation Abstracts International,* 1975, *35,* 4995A-4996A. (University Microfilms No. 74-23, 1978)

Smith, B. D. Do we need differential diagnosis at the college level? No. *Journal of Reading,* 1977, *21,* 62-66.

EDWARD FRY
MARIANNE REYNOLDS
Rutgers University

A PASSAGE-DEPENDENT TEST TO PROVIDE A DIFFERENT VIEW OF READING COMPREHENSION

Why do we need a passage-dependent reading comprehension test? We need a passage-dependent test for the following reasons:

1. *Logic.* Though a good definition of reading comprehension is lacking, there is a common assumption inherent in most reading comprehension tests that the student's ability to answer questions after reading a passage is an indication of his comprehension of that passage. This study, as well as citations of previous studies, will show that it is not necessary to read the passage in order to answer a good many questions on many well known reading comprehension tests. In other words, major reading comprehension tests have questions which are passage *independent.* This study will show that it is possible to develop a reading comprehension test which is passage *dependent* and hence meets the basic assumption that answering the questions correctly means that the information was obtained from the passage.

2. *Lower IQ Correlation.* It is fairly well known and reaffirmed in this study that major reading comprehension tests, which are passage *independent,* have a fairly high correlation with IQ tests. In other words, they tend to test the same thing. The data in this study will show that a passage-dependent test has a lower correlation with IQ than a major reading test. This has interesting implications that reading comprehension as measured on a passage-dependent test is more of an independent factor. Reading comprehension now may be more of a teachable factor, and it may be less dependent on socioeconomic status.

3. *Gain Scores.* A further reason for using a passage-dependent test is that it is superior for obtaining gain scores. The data in this paper are part of a larger study of gain scores conducted by Reynolds (1979). Most of her interesting findings will not be reported here, but it is worth noting that a passage-dependent test is far superior to a major traditional test for obtaining gain scores. This is true partly because it is difficult to get a good range of gain scores when students can answer too many questions before reading the passage.

Design of This Study

This study involved 161 eleventh and twelfth graders (eight English classes), which was nearly the total population of a suburban high school. Approximately half of them took the *California Achievement Test: Reading Comprehension,* 1977 edition, in the traditional manner; and approximately half of them took the *Passage-Dependent Test* in the traditional manner. All had previously taken The *Kuhlmann-Anderson IQ Test* in the ninth grade. Each half also took the other reading test in a gain-score mode: questions only, followed by reading the passage, followed by the questions again.

The Passage-Dependent Test

The *Passage-Dependent Test*, hereafter referred to as the PDT, was a 2100-word passage about a mythical problem situation of building a new city from the ground up in southern Alaska, written by the senior author. It was followed by 25 four-choice multiple-choice questions.

Passage Dependency

When the PDT was administered with questions only (before the students read the passage), the mean raw score was 7.48, or 30%. This was only about one item more than a chance expectancy raw score of 6.25, or 25% (S.D. 2.06).

When the *California Achievement Test*, hereafter referred to as the CAT, was administered with questions only, the mean raw score was 20.33, or approximately 50%. This was approximately 10 items, or 25% above a chance expectancy raw score of 10, or 25% (S.D. 5.64). The top score was 31, which had a Grade Equivalent of 12.9+.

When administered in a traditional fashion (namely, the students read the passage then answered the questions), the mean raw score on the PDT was 18.80, or 75%; and the mean raw score of the CAT was 26.40, or 66%. Hence, we can see that the PDT has a rather substantial gain of 45% while the traditional test had a gain of a much more modest 16%. The effect of reading the passage on the PDT was nearly three times greater in percentage points than the effect of reading the passage on the CAT. This was also a statistically significant difference ($t = 6.67$).

Relationship to IQ

The Pearson Product Moment Correlation between the CAP and the Kuhlmann-Anderson IQ scores was .70. The correlation between the PDT and IQ was .47. This is perhaps the most interesting and provocative finding of this study. Apparently a passage-dependent reading comprehension test, at least this one, correlates significantly lower with IQ than the more passage-independent CAT (Fisher $z = 2.01$).

Other Characteristics of the PDT

The reliability of the PDT was .87. The reliability of the CAT in this study was .90. However, theoretically (Spearman Brown prophecy formula), if the PDT were to be extended to 40 items, its reliability would be .91.

The PDT has approximately the same difficulty level as the CAT. The mean raw score for the PDT administered to this group of eleventh and twelfth graders was 18.8 (75%) with a standard deviation of 5.3. This compares with a mean raw score of 26.4 (66%) for the CAT with a standard deviation of 8.1 and a mean Grade Equivalent of 11.5. These scores are for both tests administered in the traditional way.

The mean IQ for the total group was 106 with a standard deviation of 14.18.

Because of the design of this particular study, it was not possible to get a direct comparison of the PDT and the CAT. It was, however, possible to get an indication by comparing the traditional administration with post-test scores. Traditional administration means the students read the passage then answered the questions. Post-test scores means that the students answered the questions, then read the passage, then answered the questions. This second answering of the questions is the post-test score. It was necessary to design the study so that we could get gain scores; hence; nobody took both CAT and PDT the traditional way. However, PDT traditional and CAT post-test correlated .76; and CAT traditional and PDT post-test correlated .63. Hence, we can state with a fair degree of confidence that there is a moderate correlation between the CAT and the PDT.

Another interesting characteristic of the PDT is that the PDT post-test score correlated with the PDT gain score .89. This high correlation means that the post-test score does essentially the same thing as the gain score; hence much time could be saved by merely giving the PDT in the traditional way (without a pre-test and calculation of gain). The probable reason for this high correlation is that in a passage-dependent test, most of the correct answers represent information gained from reading the passage. The CAT post-test score correlated with CAT gain scores a more moderate .68.

DISCUSSION

The finding of this study that a major standardized reading test has many passage-independent items is well supported in the literature. Tuinman in 1970 found that it was possible to get 34% of the items correct without reading the passage when the STEP was administered to seventh to ninth graders. In a larger study with 1200 students in grades 4 to 6, Tuinman (1973-74) found a correct responding rate ranging from 34% to 50% on the *Nelson Reading Test*, the *California Reading Test*, the *SRA Achievement Test*, the *Metropolitan Achievement Test*, and the *Iowa Test of Basic Skills.*

Passage independence has also been found by Preston (1964) for college level; Allington, Chodos, Domaracki and Truex (1977), for oral reading tests; Applebee (1972), for dull sixth graders; and Johns (1978), in grades 4 and 5.

Several studies have looked at gain scores as a measure of reading comprehension (Bloomer & Heitzman, 1965; Rankin & Lowell, 1975; Eurich, 1931), but these studies were not particularly concerned with a passage-dependent test. Incidentally, the cloze technique has not been found very effective in eliciting gain scores (Carver, 1970; Coleman & Miller, 1968; Bloomer & Heitzman, 1965) because of the small gain between pre- and post-testing.

We believe that the data show that the *Passage-Dependent Test* is an effective tool in studying information gain.

Furthermore, information gain is really what reading comprehension testing is all about. What a teacher, a principal, a parent, or the student himself wants to know is, "Was the passage understood?" If the student can answer many of the questions without even reading the passage, there is no way on earth to know if the passage was understood.

From an educational standpoint, reading teachers are perhaps being asked to do an impossible job, namely get high scores on a traditional reading test with inner-city children. It might be possible for a reading teacher to teach a student to gain information from a written passage, but it is not reasonable to ask a reading teacher to cure years of general cultural deficit. It appears that the correct answering of many questions on our major standardized reading tests are heavily influenced by a student's general knowledge. A passage-dependent test just might give a little purer measure of what might rightly be called *reading comprehension ability.*

For those who like to try to explain reading comprehension from the standpoint of schema theory, the differences between the two tests—both the differences in gain scores and the differences in correlations with IQ—might be partially explained by stating that the usual kind of reading tests like the CAT tend to test a student's pre-existing schemata rather more than new information gained from reading a passage. This would suggest that the way to improve a student's reading comprehension would be to improve his store of schema. However, this study indicates that there is also a reading comprehension skill that is perhaps partially independent of typical pre-existing schema.

For those who look at the reading world through the top-down/bottom-up theory, the PDT would seem to be more of a bottom-up type test while the CAT is more of a top-down type of test. However, these conjectures get us further and further away from the data. What we really have here is a tool, the PDT, and a bit of data which suggest that the tool is somewhat unique.

Different reading comprehension theorists, different reading test makers, different reading curriculum authors may see fit to use this tool and data as part of different, larger systems.

Having a reading test which correlates lower with IQ and meets the basic logical assumption of understanding a passage would seem to be highly desirable.

REFERENCES

Allington, R. L., Chodos, L., Domaracki, J., & Truex, S. Passage dependency: Four diagnostic oral reading tests. *The Reading Teacher,* 1977, *30* (4), 369-375.

Applebee, A. Silent reading tests: What do they measure? *Education Digest,* 1972, *37* (6), 45-47.

Bloomer, R. H., & Heitzman, A. J. Pre-testing and the efficiency of paragraph reading. *Journal of Reading,* 1965, *8* (6), 219-223.

Carver, R. P., & Darby, C. A., Jr. Analyses of the chunked reading test and reading comprehension. *Journal of Reading Behavior,* 1972-1973, *5* (4), 282-296.

Coleman, E. B., & Miller, G. R. A measure of information gained during prose learning. *Reading Research Quarterly,* 1968, *3* (3), 369-386.

Eurich, A. C. A method for measuring retention in reading. *Journal of Educational Research,* 1931, *24,* 202-208.

Johns, J. L. Do comprehension items really test reading? Sometimes! *Journal of Reading,* 1978, *21* (7), 615-619.

Preston, R. C. Responses before reading. *Journal of Educational Research,* 1964, *58,* 181-183.

Rankin, E. F., & Eberwein, L. A critical evaluation of several methods for measuring gain in reading. In G. H. McNich & W. D. Miller (Eds.), *Reading: Convention and inquiry.* Twenty-fourth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1975.

Reynolds, M. *Information gain as a measure of reading comprehension.* Unpublished doctoral dissertation, Rutgers University, New Brunswick, N. J., 1979.

Tuinman, J. Asking reading dependent questions. *Journal of Reading,* 1971, *14* (5), 289-292, 336.

Tuinman, J. Determining the passage dependency of comprehension questions in five major tests. *Reading Research Quarterly,* 1973-1974, *9* (2), 206-223.

MARJORIE HANSON
WILLIAM H. RUPLEY
JAMES F. McNAMARA
Texas A&M University

STATISTICAL MODELS AND PRACTICAL SIGNIFICANCE IN READING RESEARCH

Two kinds of statistical tests of particular interest to researchers are those that test for the significance of *relationships* and those that test for the significance of *differences*. To determine whether there are significant differences among two or more groups on a particular criterion variable outcome measure of interest, researchers frequently employ a univariate analysis of variance. This statistical model and its unique set of decision rules allows one to specify *directly* the statistical significance associated with the test of an experimental hypothesis of interest. Practical significance, on the other hand, depends on an accurate estimate of the strength of a statistical association. The practical significance assessment usually follows the design employed in tests for the significance of relationships, and often begins by asking "How much of the variance in a criterion measure can be accounted for by a prediction measure?"

Both statistical and practical significance are based on information assembled for the traditional *F*-tests employed in ANOVA designs (McNamara, 1979). However, there has been a tendency among educational researchers to base the analysis and interpretation of data entirely on the results of tests for the significance of differences (McNamara, 1978; McNamara & Gill, 1979). Overlooking practical significance, that is, failing to assess the strength of relationships, *not only* limits the information value of individual studies or experiments, *but also* constrains opportunities to integrate or synthesize empirical findings as a means to advance our theories regarding learning and instruction.

Practical significance should in no way be interpreted as a replacement for *F* values. Many research investigations test for subtle, yet important effects that may have low estimates of strengths of statistical association. However, by reporting practical significance in addition to statistical significance researchers provide their readers with information of equal importance.

The intents of this inquiry were to classify the statistical models employed in research studies reported in the *26th, 27th and 28th Yearbooks of the National Reading Conference* and to assess the practical significance for a particular subset of ANOVA findings.

CONTENT ANALYSIS

In an attempt to ascertain whether or not the reading research reported in the three most recent *NRC Yearbooks* addressed the practical significance of results, a content analysis of each report investigation was conducted. Table 1 presents the initial stage of this analysis, that is a summary of statistical models employed by researchers for each yearbook of interest.

Statistical Models

Table 1 clearly illustrates that some type of analysis of variance framework was used by a majority of investigators in each edition of the Yearbook. When *t* test applications are added to the ANOVA designs, it can further be seen that almost all significance testing employs parametric statistical models. For example, 38 of the 43 statistical models identified for the 1977 Yearbook were parametric as opposed to nonparametric frameworks.

TABLE 1

Frequency of Statistical Models Employed in the 1977, 1978, and 1979 Yearbooks of the National Reading Conference

Types of Statistical Models	Number Reported*			
	1977	1978	1979	Total
t-tests (both independent and correlated	7	10	5	22
Analysis of Variance	21	31	20	72
single level univariate	10	4	7	21
factorial univariate, analysis of covariance, multivariate	11	27	13	51
Non-parametric	5	3	3	11
Multiple Regression	2	3	5	10
Factor Analysis	1	1	2	4
Pearson Product r	7	3	7	17
Total empirical studies	33	33	33	99
Total studies	57	49	55	161

*The total number of models applied does not equal the number of studies with empirical tests; several researchers utilized more than one type of test in a single study.

The Table 1 entries are comprehensive but not complete, since a few studies were impossible to classify primarily because discussion of analyses performed was inconsistent with standard reporting procedures. Specifically, two studies both in 1977 failed to provide adequate information on supplementary parametric and nonparametric analyses and a third study (1979) reported significance levels but failed to identify the specific statistical model that produced the test results. A fourth article (1979), which could not be classified, was a well-designed synthesis of current research in a specific domain of reading. Finally, correlations that were utilized explicitly for analyses of test reliability and consistency were not entered in Table 1.

This itemized list of statistical models is useful for determining methodological trends and emphases in the field of reading research. However, the actual focus of this content analysis was to determine the amount of attention given to discussion of the strengths of relationships among factors that are found to be significant. Although ANOVA frameworks can also be formulated as multiple linear regression models (Cohen, 1968) this contemporary approach to analysis was rarely employed during this three year period. If this option had been chosen for reading research designs, consideration of the practical significance question — the amount of criterion variance accounted for by the ANOVA treatment variables — would have been an integral part of the model output.

Practical Significance

Few discussions of strength of relationship, or responses to the practical significance question, can be found in research reported in the three most recent *NRC Yearbooks.* Among the majority of studies that utilized analysis of variance designs, only four (1 in 1977, 1 in 1978, and 2 in 1979) discussed the variance accounted for by the independent variables in their designs. One nonparametric study in 1977 reported a single Cramer's V to reflect the relationship between two variables used in a Chi-square test. A study utilizing factor analysis in 1977 also reported the discriminant analysis as a test for the importance of findings derived from the factor analytic model. Although the frequently reported Pearson product moment correlation lends itself to easy consideration of strength of relationships, the r-squared value which readily serves this purpose was seldom calculated or discussed in correlational studies.

Multiple regression models contain considerations of strength of relationships as an integral part of analysis and, consequently, if any statistical results are reported from regression analysis, R^2 coefficients are included. This was true of the multiple regression models reported in the yearbooks; all studies utilizing multiple regression that reported results also included statistical measures of the strength of relationship (R^2).

Conversely, the analysis of variance models that dominated the empirical studies and also formed the focus of this content analysis rarely took the opportunity to report a related measure of practical significance. However, such a measure can easily be calculated from the statistics conventionally provided in analysis of variance studies. To illustrate, η^2, one widely accepted measure of practical significance is calculated as follows,

$$\eta^2 = \frac{F\,[df(n)]}{F\,[df(n)] + df(d)}$$

where F is the reported value for the relevant ANOVA test of significance; df(n) is the numerator degrees of freedom for the *F*-statistic; df(d) is the denominator degrees of freedom for the *F*-statistic, and η^2 is an indicator of practical significance. Eta-squared can also be viewed as the ratio of the appropriate sum of squares for treatments to the total sum of squares for a given ANOVA model (McNeil, Kelly & McNeil, 1975). Alternate practical significance indices that have been introduced in behavioral science literature include those advanced by Golding (1975), Halderson and Glasnap (1972) and McNamara (1978, 1979).

Since accepted reporting of analysis of variance design normally includes the data requirements specified in the η^2 formula offered above, were able to we calculate this practical significance index for most single factor univariate analyses reported in the three yearbooks. The results of these straightforward calculations are recorded in Table 2. It can be seen from the table that, while the majority (52) of measures of practical significance that can be calculated indicate that between 10% and 40% of variance in the dependent variable can be explained by the independent variable, a substantial number (24) indicate less than 10% of variance explained. A few (11) involve independent variables that account for 40% or more of the variance in the dependent measure. In view of the noticeably high η^2 values obtained from the Yearbook studies, it is somewhat puzzling that researchers in reading are not more open about reporting them, as they contribute to the strength of findings.

TABLE 2

Indicators of Practical Significance for Single Level Univariate Models* Reporting Statistical Significance

Eta Squared Ranges		Yearbooks 77	78	79	77-79
.0100	.0499	2	3	1	6
.0500	.0999	11	6	1	18
.1000	.1999	8	6	2	16
.2000	.2999	9	4	10	23
.3000	.3999	6	3	4	13
.4000	.4999	1	–	2	3
.5000	.5999	1	1	2	4
.6000	.6999	–	1	1	2
.7000	.7999	–	1	1	2
Number of tests		38	25	24	88
Number of studies (with sufficient information		10	8	8	27
Number of studies (with insufficient information)		4	2	1	7

* t-tests of significance are assumed to be independent and thus equivalent to a single-level dichotomous F in the calculation of these eta^{2}s. One series of t-tests (1979) which are clearly identified as correlated is omitted from the calculations as well as post-hoc tests of significance in ANOVA designs.

The question inevitably arises as to what is an "acceptable" level of practical significance. Few authors have ventured to specify a cut-off point that would determine arbitrarily at what point a statistically significant finding becomes practically significant. Applying such a criterion to all aspects of educational research could preclude discussion of findings low in practical significance that are nevertheless of some import. Additionally an arbitrary criterion overlooks contextual factors in designs, such as measurement and sampling, which also contribute to the value of researched information. Therefore, the level of practical significance deemed valuable must be left to the judgment of the researcher just as the level of statistical significance is becoming a matter of judgment (Rupley, 1977; Rupley & McNamara, 1979).

Cohen (1977) suggests the following operational guidelines for interpreting the practical significance of behavioral science data in terms of correlation coefficients. "Small" effects are characterized by an $r = 0.10$; "medium" effects are indicated when an $r = 0.30$; and "large" effects are evident when $r = 0.50$. Hence, Cohen's position implies that a large effect would be encountered when at least 25 percent of the variance is attributable to the independent variable.

Adams et al. (1970) selected a 0.10 cut-off point for findings to be of practical signficance in their nationwide study. Using their criterion, 63 out of 87 of the tests for which practical significance is calculated in our analysis of the *NRC Yearbooks* would be significant. It would seem logical, however, to pursue for further study and replication the studies that report findings in which the greatest proportion of variance is explained in hopes of discovering factors and methodologies in the practice of reading instruction that have the greatest impact on student learning. Thus, the relative amount of practical significance, rather than an arbitrary level such as 10% or 20%, might be most important to note in a field as prolific as reading research is today.

Based on the results of our intensive content analysis, the authors offered the following for careful consideration.

(1) The η^2 statistic or some other measure of practical significance would be useful in directing reading researchers' attention to those variables that warrant further study in subsequent investigations.

(2) Research reported in the *NRC Yearbook* frequently includes insufficient data to enable an assessment of the practical significance of results. Descriptions of designs were often imbedded in the text and not sufficiently clear to enable the reader to reconstruct the complete ANOVA table.

(3) Furthermore, even in those studies where we were able to calculate the η^2 statistics, there was often some question as to the accuracy of the degrees of freedom. In these cases, if the description of the sample and of the treatment variables had been comprehensive, the model could have been reconstructed and the correct degrees of freedom ascertained.

(4) A need for more accurate descriptions of sampling plans was evident since sample selection is likely to have a decided influence on both tests of differences and tests of relationships. Furthermore, discussion of generalizability especially vital in view of the extremely small and nonrandom samples which were typically utilized was notably absent.

(5) In several of the ANOVA studies a series of univariate tests were performed utilizing the same independent variables to assess significant differences for several distinct criterion variables without any reference to the possibility that one or more of the resulting individual F test differences might be due merely to chance since the set of F tests are not statistically independent. In such cases a single multivariate design would have been more appropriate.

(6) Nonparametric techniques were seldom used, although the characteristics of the researcher-developed instruments utilized in some of the studies suggests that the measures would fail to meet at least the assumption of continuity in the dependent variable.

IMPLICATIONS FOR READING RESEARCHERS

While our treatment of practical significance in reading research focused primarily on the use of η^2 to aid in the interpretation of the importance and magnitude of differences identified in analysis of variance tests of significance, it is not implied that attention to practical significance represents a concern just recently entered into the educational literature. Practical significance in educational research was initially addressed in the mid 1960s (Schutz, 1966), but research reported in education journals has persisted in neglecting discussion of this important issue (McNamara, 1979). The content analysis reported in this inquiry points out that in 1977 through 1979 — over a decade later — attention to practical significance has not yet found its way into the reading research published in the *NRC Yearbooks*.

Reporting Research

Most statisticians agree that η^2 (or a comparable index of practical significance) communicates a great deal more information than the traditional F value. It takes into account sample size and thus provides a more accurate representation of the degree of prediction. Furthermore, since η^2 is calculated directly from ANOVA results, it creates no difficulty for the investigator.

Given the fact that considerable interest is presently being given to meta analysis and secondary analysis of educational research, reading researchers' reporting of practical significance needs to be given paramount consideration. Athough one cannot accurately specify how often the research reported in any given reading journal is consulted or cited as a reference for various decision making situations, authors should report an index of practical significance. By putting their readers on notice as to their position on practical significance requirements, researchers no longer restrict the users of such research to merely referencing or counting the number of statistically significant differences on criterion variables of interest.

It is our contention that reading researchers should report η^2 (or a comparable index) in journal articles, reviews of research, technical reports and dissertations. Also, practical significance values should be in the forefront when any reading researcher finishes the results section of his or her study and enters the section on discussion of findings.

Training Researchers

To demonstrate that the η^2 value communicates a good deal of information, and thus provides reading researchers more accurate measures of the relative reduction of uncertainty that can be directly linked with a particular research inquiry, professors who train future researchers or program evaluators might initiate one or more of the following alternative activities.

Alternative One. In seminars devoted to reading research design, an interesting project might be to conduct a practical significance analysis on all recently completed doctoral dissertations. When classification errors (low practical significance values for highly significant difference tests) are encountered, students could be asked to prepare revised interpretations and discussions of dissertation results. Comparisons of original and revised versions provided excellent means to illustrate the consequences of viewing tests of significance on an experiment as the ultimate objective.

Alternative Two. In seminars devoted to theory development or knowledge synthesis, professors should encourage students to compute and include η^2 values (or their equivalent) for each research study reviewed. In the past, probability values (i.e., $p \leqslant 0.01$) have generally been used to support theories. The intent here would be to suggest that an index of practical significance in conjunction with probability values should be used to indicate the satisfaction students have with their theoretical frameworks.

Alternative Three. If professors are able to locate current operational studies where the probability values associated with mean difference ANOVA tests were "highly significant" but the η^2 values were "extremely low," they might use a seminar group to conduct an organizational analysis (case study) to determine how administrators and other policy-making groups in the organization actually use their "significant differences" tests results in subsequent decisions (i.e., reporting results to the public, creating new priority issues, preparing press releases, changing teacher assignments, or implementing new instructional designs). Analysis of an *actual* organizational example and the *real* consequences that are uncovered in that setting may be far more convincing to graduate students whose immediate professional interests center primarily on practice rather than theory development and verification.

SUMMARY

The implications suggested above are intended to give reading researchers the confidence to note the following: First, when "highly significant" ANOVA test results are accompanied by "extremely low" eta squared values, one would do well to follow the advice of McNeil and his colleagues (1975) who interpret this situation as "knowing something for sure about very little." Second, when high η^2 values are continually uncovered on the same instructional variables, the field should be encouraged to recognize their potential value for the development of theory in reading research.

Hopefully, awareness of these facts will encourage the reporting of practical significance in future reading research investigations.

ACKNOWLEDGEMENT

Preparation of this article was supported in part by the Reading/Language Research Laboratory and in part by the Instructional Research Laboratory, Department of EdCI, Texas A&M University, College Station, TX.

REFERENCES

Adams, R. S., Kimble, R. M., & Martin, M. School size, organizational structure and teaching practices. *Educational Administration Quarterly*, 1980, *6* (3) 15-31.

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, *70*, 426-443.

Cohen, J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1977.

Golding, S. L. Flies in the ointment: Methodological problems in the analysis of the percent of variance due to persons and situations. *Psychological Bulletin*, 1975, *82*, 278-288.

Halderson, J. S. & Glasnap, D. R. Generalized rules for calculating the magnitude of an effect in factorial and repeated measures ANOVA designs. *American Educational Research Journal*, 1972, *9*, 301-310.

Jackson, G. B. *Methods for reviewing and integrating research in the social sciences.* Final report, National Science Foundation Grant # DIS 76-20398. Washington, D.C.: Social Research Group, George Washington University, April 1978.

McNamara, J. F. Practical significance and statistical models. *Educational Administration Quarterly,* 1978, *14* (1), 31-50.

McNamara, J. F. Practical significance and mathematical models in administrative research. In G. L. Immegart & W. L. Byrds (Eds.), *Problem-finding in educational administration: Trends in theory and research.* Toronto: Lexington Books, 1979.

McNamara, J. F. & Gill, D. R. Practical significance in vocational education research. *Journal of Vocational Education Research,* 1979, *3* (1), 20-38.

McNeil, K. A., Kelly, F. J. & McNeil, J.T. *Testing research hypotheses using multiple linear regression.* Carbondale, Ill.: Southern Illinois University Press, 1975.

Rupley, W. H. Teacher instructional emphasis and pupil achievement in reading. *Peabody Journal of Education,* 1977, *54,* 286-291.

Rupley, W. H. & McNamara, J. F. Effects of instructional emphases and pupil engaged time reading achievement. In M. L. Kamil & A. J. Moe (Eds.), *Reading research: Studies and applications.* Twenty-eighth Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1979.

Schutz, R. E. Research in schools. *Educational Researcher,* 1966, *5,* 151-158.

MARGARET MacLEAN
McGill University

QASOR: A FRAMEWORK FOR QUALITATIVELY ANALYZING SILENT AND ORAL READING

Although the market is repleat with diagnostic instruments which purport to assess oral reading behavior, many of these instruments lack a coherent conceptual framework for error analysis. Typically, oral reading is analyzed by quantifying errors into arbitrary categories which frequently lack appropriate theoretical rationale.

One exception to this quantitative, atheoretical approach to oral reading error analysis is the *Reading Miscue Inventory* (RMI) (1972) which allows qualitative analysis within a well-defined theoretical framework. Through careful, detailed analysis of oral reading miscues, the RMI allows one to infer qualitative, as well as quantitative, information about the extent to which a reader attends to and uses language cues to extract meaning as s/he reads.

While acknowledging that miscue analysis procedures such as the RMI provide valuable insights into the reading process, some researchers have questioned the comprehensiveness, accuracy, and practicality of the procedures. Newman (1978) cautions against the exclusive focus of miscue analysis on oral reading. He contends that a more meaningful assessment of reading should include both oral and silent reading. Wixson (1979) suggests that miscue analysis procedures do not account for the influence of such factors as instructional method, level of reading proficiency, a reader's prior knowledge, and the nature of the written text, all of which research has suggested have important influences on the reading process. She indicates that "...it is possible that miscue patterns as identified by standard miscue procedures do not provide an accurate reflection of the reading process" (Wixson, 1979, p.170). She proposes that further refinement of miscue analysis procedures, taking into account some of the above variables, would seem appropriate. Hood (1978) contends that the complex and time-consuming procedures for analyzing and interpreting miscues limits the practicality of the procedures.

In view of these cautions regarding standard miscue procedures, it may be time to consider developing alternative diagnostic procedures. The Qualitative Analysis of Silent and Oral Reading (QASOR) (Aulls & MacLean, 1978) is proposed as an alternative framework for investigating individual variations in oral and silent reading. Based on similar psycholinguistic principles as the RMI, QASOR proposes

1. including measures of both oral and silent reading, using miscue analysis and cloze procedures, on both narrative and expository passages.
2. including an estimate of prior knowledge of key concepts relevant to the passages read, using the PReP procedure (Langer, 1979).
3. focusing on meaning acquisition through the examination of the semantic acceptability of miscues, clozures, and self-corrections, at both the sentence and text levels.

4. focusing on syntactic, semantic, and graphophonic integration through examination of miscues, clozures, and self-corrections.
5. examining qualitative differences' in miscues/clozures by sampling from the beginning and final portions of text.

By including measures of oral and silent reading on both narrative and expository passages, by assessing prior knowledge, by analyzing self-corrections as well as miscues, and by comparing miscues/clozures from different portions of text, QASOR provides a framework for examining how well a reader constructs and reconstructs meaning as s/he reads through a text.

RATIONALE AND PROCEDURES

Oral and Silent Reading of Narrative and Expository Text

It is important not to overgeneralize oral reading performance to silent reading. Whereas oral reading emphasizes a forward moving process, during silent reading, a reader is able to re-read and re-check information. In addition, Newman (1978), suggests that silent reading allows a reader more time to use graphophonic, syntactic, and semantic cues.

QASOR proposes assessing both oral and silent reading. Oral reading is assessed through miscue analysis, while silent reading is assessed through a modified cloze procedure. Cambourne (1977) and Farren (1977) have suggested that just as miscues can be analyzed to provide information about the enroute oral reading process, so too, cloze responses can be examined to gain insights about the enroute silent reading process. By comparing error patterns in both oral and silent reading, one can estimate to what extent a reader is using similar strategies.

To obtain data on both oral and silent reading, the reader is requested to read, through a text silently, completing cloze on selected portions as s/he reads. After reading the whole text silently, s/he is then directed to read specific portions orally.

Each reader reads, silently and orally both an expository and a narrative passage. Miscue research (Brazee, 1976; Carlson, 1970) has indicated there are qualitative differences in miscues on narrative and expository texts. As well, Wixson (1979) contends that the structure and content of reading materials influences miscue patterns. Performance on one type of text may not be generalizable to other text types. As classroom reading material typically includes both narrative and expository text, assessment of reading behavior should include both types of material. Admittedly, there are a variety of different types of narration and exposition. Following QASOR guidelines, one should try to select narratives which follow the general organizational structure characteristic of typical narratives, that is, setting, theme, plot, resolution (Thorndyke, 1977), and expository selections which are self-contained and develop only one major topic. As well, subtopics should clearly refer to each other and the main topic. It is important that passages be well-organized and cohesive, in order that a reader can take advantage of the redundancy and referential structure within the passage.

It is important too, that passages be of sufficient length, as miscue research (Menosky, 1970) has indicated that the quality of miscues changes when short passages are read. Goodman and Goodman (1977) suggested that short passages do not provide the reader with sufficient cues to develop a sense of the author's meaning and style. "Longer passages offer redundant opportunities to recover and self correct" (Goodman & Goodman, 1977, p. 330). Smith (1971) contends that longer passages are easier to read than shorter ones as they usually provide more extensive context and greater redundancy, and therefore, more opportunities for predicting and acquiring meaning.

The QASOR procedures recommend using texts approximately 1200 words or longer. This length more closely reflects the average length of natural text than the short passages typically used in standardized and informal tests. As well, this length allows an adequate number of cloze deletions to make confident inferences about a reader's use of semantic and syntactic redundancy to predict deleted words.

Prior Knowledge Assessment

Of itself, print carries no meaning. Rather, it is the reader's prior knowledge which allows the individual to anticipate and interpret meaningfully an author's intended message. For a reader to process text successfully, new concepts and vocabulary must be linked to prior knowledge.

Miscue research provides evidence that prior knowledge does influence miscue patterns. Thomas (1975) found that a reader's tendency to produce syntactically and/or semantically acceptable miscues depends to some extent on the degree to which s/he understands the material. Rousch (1972) found that average grade four readers who lacked prior knowledge of material being read, generated more graphophonic miscues and made fewer self corrections than readers with more complete prior knowledge.

QASOR proposes including an assessment of prior knowledge using the PReP procedure developed by Langer (1979). This procedure attempts to scale quality of knowledge (from little to much) that a reader possesses about a concept, in terms of the extent to which relevant aspects of its meaning are organized.

Using PReP, one examines the text to be read to identify one or more dominant concepts which are essential to the cohesion of the text structure. The reader is requested to provide as much information as possible about the concept(s) selected. This information is then scored according to PReP guidelines to determine the extent of the reader's prior knowledge for the passage. This prior knowledge score can be compared with miscue patterns to determine the extent to which the two interrelate.

Meaning Acquisition

QASOR suggests examining the extent to which a reader attends to meaning by examining the semantic acceptability of miscues, clozures and self-corrections. By examining semantic acceptability at both the sentence and text level, one can attempt to determine the extent to which the reader is adding up information as s/he reads through a text. By examining self corrections, one can obtain another estimate of a reader's awareness of semantic information. Good readers tend to self correct selectively, whereas, poor readers either do not self correct, or they tend to over correct. By examining the semantic acceptability of miscues and clozures, and by assessing the proportion of meaningful self corrections, one can estimate to what extent a reader is acquiring meaning as s/he processes text.

Integration of Syntactic, Semantic, and Graphophonic Information

Like other miscue analysis procedures, QASOR recommends examining miscues and self corrections to determine whether they are semantically acceptable within the sentence or the context of the text, whether they are syntactically acceptable within the sentence, and/or, if they are graphophonically similar to the text word, in at least two of three parts. Unlike standard miscue procedures however, QASOR extends this examination of syntactic, semantic, and graphophonic acceptability to cloze replacements. Miscues and self corrections at both the text and sentence levels are considered, in order to determine the extent to which a reader is able to integrate the information processed during oral and silent reading.

Progressive Sampling of Oral and Silent Reading

Menosky (1971), in her analysis of oral reading miscues across varying portions of text, found that depending on a reader's proficiency, the quality of miscues changed across text. Poor readers produced more unacceptable miscues in the beginning than in subsequent portions of text.

QASOR recommends sampling oral and silent reading from the beginning and final portions of a text to determine if the reader is benefitting from cumulative textual information. This progressive sampling of reading can be achieved by dividing a text into equal segments (approximately 200 words each). Cloze passages are prepared on selected segments near the beginning and end of the text, for a *minimum* of fifty deletions. (If a text is divided into six equal segments, cloze could be prepared on segments three and six.) A reader is directed to read orally intact beginning and final segments (for example, segments two and five), after reading the whole text silently, completing the selected close portions. Oral and silent reading performance on beginning and final portions of text can then be compared to see if there are differences in the quantity and quality of miscues.

PROPOSED QUALITATIVE INSIGHTS

As it includes parallel measures of oral and silent reading, on both narrative and expository text, potentially, at least, the QASOR framework appears to offer diagnostic insights relevant to the classroom teacher. These insights can be considered in terms of the following six questions.

1. To what extent does the reader demonstrate similar strategies in oral and silent reading?

That is, does the fact that oral reading emphasizes a forward moving process, whereas during silent reading a reader can reread and recheck information, influence the quantity and quality of miscues, clozures, and self corrections? Poor readers who have many nonmeaningful miscues, will likely exhibit a similar proportion of miscues/clozures. This happens because they have not learned strategies for attending to syntactic and semantic information, but focus primarily on graphophonic information. In contrast, relatively fluent readers may vary considerably in the proportion of meaningful miscues exhibited in oral and silent reading. Frequently these differences will reflect a higher meaning acquisition score in silent, than in oral, reading. This pattern indicates that the reader is treating reading as both a construction and reconstruction process. S/he is able to compensate for within sentence prediction errors during silent reading by selectively reconstructing those miscues which are incongruent with the overall contextual meaning.

2. To what extent does the reader demonstrate similar strategies on narrative and expository text?

By comparing the quantity and quality of miscues, clozures, and self corrections, on both text genres, one can estimate whether a reader is able to process both with equal facility. If the pattern of miscues differs significantly between these text types, it could be due to the fact that the reader lacks appropriate strategies for reading different text structures. Facility with narrative does not necessarily mean a reader will be able to cope effectively and efficiently with expository text. By comparing performance on both narrative and expository passages one can more confidently generalize how well a student will deal with classroom material which typically includes both.

3. To what extent does the reader use prior knowledge to facilitate text comprehension?

Does a reader with a high prior knowledge score produce better quality miscues than the reader with a low prior knowledge score? By comparing a reader's prior knowledge score with the pattern of miscues, one can estimate how well s/he was able to use prior knowledge to process the text. If a reader demonstrates minimal knowledge about key concepts developed in a text, one would anticipate that s/he would have difficulty predicting and interpreting the text. Therefore, the proportion of nonmeaningful miscues would likely be high. As well, the reader would not likely make many meaningful self-corrections. By including an assessment of prior knowledge one can estimate the extent to which miscues are related to a lack of conceptual background, or due to confusions in processing syntactic, semantic, and graphophonic information.

4. To what extent does the reader attend to semantic information in the text?

By examining the proportion of miscues/clozures that make sense, one can estimate the extent of a reader's awareness that reading should be a meaning-centered process. A reader whose meaningful miscues exceed his/her miscues is demonstrating that s/he knows that reading should sound like language and make sense. By comparing the ratio of meaningful to non-meaningful miscues/clozures, one can estimate how well a reader gains and uses meaning.

Examination of self corrections can provide additional information about meaning acquisition. A relatively fluent reader's self corrections will likely be sensible at either the sentence or total text level. This reader tends to self correct selectively, correcting only those miscues which interfere with meaning. A reader with fewer meaningful than non-meaningful miscues will likely overcorrect or not correct at all. Overcorrectors tend to ignore meaning by focusing on the most immediate information being processed at the expense of integrating it with the larger meaning of the text message. Readers who make few or no self corrections seem not to perceive reading as involving the construction and reconstruction of meaning. Reconstruction involves correcting miscues on the basis of information that precedes, as well as follows.

5. To what extent does the reader integrate syntactic, semantic, and graphophonic information?

By considering whether a reader's miscues/clozures and self corrections are syntactically, semantically, and graphophonically acceptable, one can determine how well the reader integrates linguistic information. The proficient reader, more likely than the less proficient, produces miscues, clozures and self corrections which are syntactically and semantically acceptable. By comparing miscues and clozures one can assess to what extent the reader uses similar strategies during oral and silent reading for processing text.

6. To what extent does the reader use text redundancy?

By comparing the quantity and quality of self corrections, miscues, and clozures from the beginning and final portions of text, one can attempt to determine if the reader is benefitting from cumulative information acquired as s/he reads through a text. Fluent readers, those who are able to take full advantage of semantic redundancy and textual connectiveness, will likely make fewer miscues in the final than beginning portions of text. Less fluent readers, however, those who are less able to use the full range of text redundancy, will likely make a similar proportion of miscues in the beginning and final portions of text. The reader with a higher proportion of meaningful than nonmeaningful miscues/clozures in the final portions of text is assumed to be more capable of integrating text information.

These six proposed qualitative insights provide some indication of the qualitative information which QASOR attempts to assess. Considering aspects of both reader (prior knowledge) and text (narrative and expository texts), QASOR proposes examining how these factors interrelate to produce various miscue patterns in oral and silent reading. The qualitative analysis and comparison of oral and silent reading performance allows one to have more than one window on the reading process. Although still a proposed framework, rather than a diagnostic instrument, QASOR appears to represent a viable, qualitative alternative to skills-focused quantitative assessment devices.

REFERENCES

Aulls, M., & MacLean, M. *QASOR: Qualitative analysis of oral and silent reading behaviors.* Unpublished manuscript, 1978.

Brazee, P. *A qualitative and quantitative description of eighth grade students' oral reading in both narrative and expository materials.* Unpublished doctoral dissertation, University of Northern Colorado, 1976.

Cambourne, B. *Some psycholinguistic dimensions of the silent reading process: A pilot study.* Paper presented at the Annual meeting of the Australian Reading Conference, Melbourne, 1977.

Carlson, K. L. *A psycholinguistic description of selected fourth grade children reading a variety of contextual material.* Unpublished doctoral dissertation, Wayne State University, 1970.

Farren, S. *A study of reading errors using Goodman's miscue analysis and cloze procedure.* Paper presented at the Annual meeting of the Reading Association of Ireland, Dublin, 1977.

Goodman, K., & Goodman, Y. Learning about psycholinguistic processes by analyzing oral reading. *Harvard Educational Review*, 1977, *47*, 317-333.

Goodman, Y., & Burke, C. *Reading Miscue Inventory: Manual.* New York: Macmillan, 1972.

Hood, J. Is miscue analysis practical for teachers? *Reading Teacher*, 1978, *32* (3), 260-266.

Langer, J. *Pre-reading plan (PReP): How to make a hard text easier to read.* Unpublished manuscript, 1979.

Menosky, D. *A psycholinguistic analysis of oral reading miscues generated during the reading of varying portions of text by selected readers from grades 2,4,6,8: A descriptive study.* Unpublished doctoral dissertation. Wayne State University, 1971.

Newman, H. Oral reading miscue analysis is good but not complete. *Reading Teacher*, 1978, *31* (3), 883-885.

Rousch, P. *A psycholinguistic investigation into the relationship between prior conceptual knowledge, oral reading miscues, silent reading, and post-reading performance.* Unpublished doctoral dissertation, Wayne State University, 1972.

Smith, F. *Understanding Reading.* New York: Holt, Rinehart, & Winston, 1971.

Thomas, K. J. *Syntactic/semantic acceptability and semantic similarity of oral reading errors as functions of variation in attained comprehension.* Unpublished doctoral dissertation, University of Arizona, 1975.

Thorndyke, P. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9* (1), 77-110.

Wixson, K. Miscue analysis: A critical review. *Journal of Reading Behavior*, 1979, *11* (2), 163-175.

MARY BETH MARR
KATHLEEN R. LYON
State University of New York at Albany

PASSAGE INDEPENDENCY: AN EXAMINATION OF THREE INFORMAL READING INVENTORIES

Reading specialists and classroom teachers commonly administer informal reading inventories (IRIs) to assess a student's strengths and weaknesses in reading. These inventories are also used to identify a general level of reading proficiency for placement in appropriate instructional materials. If teachers are to have confidence in the quality of these assessments and subsequent instructional decisions, they need to be certain that the measures they are using are valid.

Informal reading inventories provide a number of estimates including sight word vocabulary, oral reading accuracy, fluency, rate and comprehension. The latter was of particular interest in this investigation. Typically reading comprehension refers to an individual's ability to understand text. In recent years considerable attention has been given to passage independency, the extent to which questions can be answered correctly without reading the passages on which they are based (Johns, 1978; Pyrczak, 1972, 1974). If test questions can be answered without reading the corresponding passages, it is argued that the test does not really measure reading comprehension, but rather prior knowledge. Thus, an adequate comprehension score may not reflect the extent to which a student has understood the passage, but merely his/her ability to remember information about the topic.

Researchers have found that several commonly used comprehension tests are in fact passage independent. Preston (1964), Tuinman (1973-1974), and Scherick and Hanna (1977) examined standardized reading achievement tests such as the *Nelson Denny Reading Skills Test* and the *Cooperative English Test* and found that students were able to answer comprehension test questions above the chance level without reading the corresponding paragraphs. Allington, Chodos, Domaracki and Truex (1977) found similar results when examining four diagnostic reading tests: the *Spache Diagnostic Reading Scales*, the *Durrell Analysis of Reading Difficulty* and the *Gilmore* and *Gray Oral Reading Tests*. Reading comprehension exercises including the *Readers Digest Skill Builder* series and the *McCall-Crabbs Standard Lessons in Reading* have also been found to be passage independent (Pyrczak & Axelrod, 1976). In light of these findings it has become increasingly apparent that questions can in fact be passage independent thus reducing their validity as a measure of comprehension. The intent of the present study was to evaluate the passage independency of questions taken from three commonly used informal reading inventories: the *Classroom Reading Inventory* (CRI) Form B, Fourth edition by Silvaroli (1978), the *Analytical Reading Inventory* (ARI) Form C, by Woods and Moe (1977) and the *Ginn 720* basal inventory (Ginn and Company, 1976). To date this type of IRI evaluation has not been undertaken. A second focus of the study was to identify passage independent question types; therefore, an analysis of student responses to specific questions was conducted.

METHOD

Subjects

Thirty-six fourth grade students from a rural school in upstate New York participated in the study. They were stratified randomly into one of three IRI test groups wherein there was an equal number of good and poor readers. Reading ability was determined by performances above or below grade level on both vocabulary and comprehension subtests of the *Iowa Test of Basic Skills.*

Procedure

All children were tested individually using a randomly selected form (A,B or C) of each inventory. Test questions were read orally to the children according to the inventory test procedures. However, children did not have access to the passages from which the questions were derived. The order for presenting questions within each level (primer through sixth grade) was maintained as in the original inventories, but the levels were presented in one of two randomized orders to avoid making the testing task one of increasing difficulty for the students. Children's responses to the questions were tape recorded and later transcribed for evaluation.

RESULTS AND DISCUSSION

Due to the unequal number of questions contained in the three inventories, the percentage of items answered correctly was determined for each test. An analysis of variance procedure was conducted to determine if these percentages were in fact significantly different from one another. The results indicated that performances across the three tests were different, $F(2,30) = 15.668$, $p \leq .001$. The CRI was found to be the most passage independent (M = 26.273) followed by the ARI (M = 17.455) and Ginn inventories (M = 11.455) respectively. Newman-Keuls post hoc contrasts revealed that each of these mean percentage scores were significantly different from one another. These results point out that although all three tests were found to be passage independent to some degree, the CRI contained the greatest number of passage independent questions (26%) and the Ginn inventory the fewest (11%). Thus, the extent to which comprehension scores could be inflated on these tests varied significantly.

An item analysis revealed two interesting types of information: 1) the number of children able to test out of a particular reading level by answering passage independent questions correclty and 2) the types of questions which were passage independent in nature.

On the comprehension subtest of the CRI, an instructional reading level cutoff score is 60% correct. An item analysis revealed that at level #1 of the test, 50% of the children passed the instructional reading level criterion without ever reading the passages. At levels #2 and #4, 25% of the students passed the criterion score. Based upon these percentages one could predict that if one hundred children were tested, fifty would test out of level #1 and twenty-five out of levels #2 and #4 without reading the passages. These findings lend considerable question as to the validity of this comprehension measure.

The number of children passing the criterion score for the ARI and Ginn inventories was considerably less. On the ARI approximately 33% of the students scored above criterion at level #2 without reading the paragraphs. However, they did not test out of the other grade levels on this inventory. On the Ginn inventory, two children passed the criterion score at level #5; while at levels #6 and #9 only one child was able to meet the criterion without reading the corresponding paragraphs.

Of the three tests examined the CRI appears to be the least valid measure of text comprehension. It not only contains the greatest number of passage independent questions which results in students passing comprehension criterion levels without ever reading the passage. It also provides the teacher with an inflated estimate of the student's comprehension ability and thus increases the probability that the student will be placed in instructional materials at an inappropriate level of difficulty.

TABLE 1

Passage Independent Test Questions

Test	Level	% Correct Response[a]	Question
CRI	P	75	How did the children feel?
CRI	1	100	Where do ants live?
CRI	1	70	What does the word "busy" mean in the story?
CRI	2	79	Why did the people run from the bulls?
CRI	2	75	What does the word "chase" mean?
CRI	4	75	What made riding in a Conestoga wagon unpleasant?
CRI	4	91	What are some ways of travel found today?
ARI	2	79	What is meant by the phrase "fight back the tears"?
Ginn 720	5	92	How do you think Mike felt when mother said no?
Ginn 720	6	92	Do the animals like living in the tree?
Ginn 720	6	83	What do you think the blackbird will do with the hat?
Ginn 720	11	75	What kind of person do you think Mr. Green is?

[a] n = 12

Note. Only questions which 70% or more of the students answered correctly were reported.

Examination of the types of questions which were passage independent reveals some definite patterns. The questions listed in Table 1 tend to fall into three general categories: 1) general information 2) vocabulary meaning 3) affective. The general information questions appear to tap information about familiar topics or prior experiences as evidenced by the questions:

Where do ants live? What are some ways of travel found today? and Why did people run from the bulls? The vocabulary questions contain words or phrases already a part of the student's oral vocabulary (e.g. busy, chase, fight back the tears). Hence a student does not have to read the passages to derive the word meanings. Questions in the affective category may elicit a word association such as that between *no* and *sad* or *feel* and *good* which could account for the number of correct responses. Samuels (1968) notes that strength of word association is often a salient feature in children's guessing behavior. Examples of questions in this category include: How did the children feel? and How do you think Mike felt when mother said no? These data suggest that the type of question asked can signal the extent to which it might be passage independent.

CONCLUSION

An evaluation of these three informal reading inventories has pointed out significant differences between and among the tests with regard to passage independence, characteristics of passage independent questions and the potential to inflate composite comprehension scores. While a passage independency analysis appears necessary in evaluating the validity of reading comprehension tests, it is not sufficient. As Hanna and Oaster (1978-1979) point out, passage independence is not the complement of passage dependence. That is, an item which cannot be answered correctly when the passage is not read may not be able to be correctly answered after reading the passage. Hence, in the construction of future comprehension tests or the revision of existing measures, authors will want to thoroughly evaluate test questions to identify both passage dependent and independent characteristics. Certainly the best way to identify these characteristics is to pilot the passages and questions on a prototypical population of students for which the test is designed. The present investigation completes the passage independency facet of this evaluation process.

REFERENCES

Allington, R.L., Chodos, L., Domaracki, J., & Truex, S. Passage dependency: Four diagnostic oral reading tests. *The Reading Teacher,* 1977, *30,* 369-375.

Ginn & Company, *Ginn 720: Basal Series.* Boston: Ginn and Company, 1976.

Hanna, G.S., & Oaster, T.R. Toward a unified theory of context dependence. *Reading Research Quarterly,* 1978-1979, *14,* 226-243.

Johns, J.L. Do comprehension items really test reading? Sometimes! *Journal of Reading,* 1978, *21,* 615-619.

Preston, R.C. Ability of students to identify correct responses before reading. *The Journal of Educational Research,* 1964, *58,* 181-183.

Pyrczak, F. Objective evaluation of the quality of multiple-choice test items designed to measure comprehension of reading passages. *Reading Research Quarterly,* 1972, *8,* 62-71.

Pyrczak, F. Passage dependence of items designed to measure the ability to identify the main ideas of paragraphs: Implications for validity. *Educational and Psychological Measurement,* 1974, *34,* 343-348.

Pyrczak, F., & Axelrod, J. Determining the passage dependence of reading comprehension exercises: A call for replications. *Journal of Reading,* 1976, *19,* 279-283.

Samuels, S.J. Effect of word associations on reading speed, recall, and guessing behavior on tests. *Journal of Educational Psychology,* 1968, *59,* 12-15.

Scherick, H.H., & Hanna, G.S. Passage dependency data in the selection of reading comprehension test items. *Educational and Psychological Measurement*, 1977, *37*, 991-997.

Silvaroli, N.J. *Classroom reading inventory*. Dubuque, Iowa: William C. Brown Company, 1976.

Tuinman, J.J. Determining the passage dependency of comprehension questions in 5 major tests. *Reading Research Quarterly*, 1973-1974, *9*, 206-233.

Woods, M.L., & Moe, A.J. *Analytical reading inventory*. Columbus, Ohio: Charles E. Merrill Publishing Company, 1977.

MICHAEL C. McKENNA
JOHN W. MILLER
Wichita State University

THE EFFECTS OF AGE AND DISTRACTOR TYPE ON MAZE PERFORMANCE

The maze procedure (Guthrie, 1973) is one of a variety of multiple-choice cloze formats. In maze, subjects reading a prose text periodically encounter vertical arrays of words,
they,
like these, only one of which is acceptable in context. The task therefore
could,
approximates an actual maze, and the correct "pathway" is to be indicated in each case by circling the right answer.

The establishment of reliable scoring criteria for the use of maze in determining reading levels is a central goal of current research. The fact that such criteria have been established for the cloze procedure is somewhat encouraging, but maze tests are inherently susceptable to sources of variance not found in cloze. These include (1) the number of distractors per item, (2) the arrangement of distractors, and (3) the nature of distractors. From a passage of 250 words, for example, only five fifth-word cloze tests are possible. However, the number of fith-word maze tests in which two distractors per item are selected from a very short list of 1000 words greatly exceeds 10^{100} (Wiseman & McKenna, 1978).

The fact that appreciable differences have been observed even among the five cloze tests over the same passage (Bormuth, 1964) suggests that the additional sources of variation in maze tests may magnify the problem. In one of the few studies of the issue to date, Bradley and Meredith (1978) concluded that maze may well be unsuitable for use at the junior high level. On the other hand, the establishment of rigorous controls in the production of maze tests has been observed to result in acceptably high alternate-form reliability (Bradley, Ackerson, & Ames, 1978). Evidence is virtually nonexistent, however, of which controls are likely to be most effective and of the principles which govern their effectiveness.

The present study investigates two potential sources of variance in maze performance, age and distractor type, while holding constant the number of distractors per item and the arrangement of the distractors. In particular, the following hypotheses are tested:

1. Do third- and sixth-grade students differ in their ability to accurately respond to maze items?
2. Is the accuracy level of responses to maze items affected by distractors varying in graphic similarity or form class?
3. Is the accuracy level of third- and sixth-grade students in responding to maze items differentially affected by distractors varying in graphic similarity or form class?

METHOD

Subjects

A total of 20 third-grade students and 20 sixth-grade students were selected on the basis of Reading Comprehension subtest scores on the *California Achievement Test* (1970 ed.). Only students reading at or above grade level were included in the study. Students reading below grade level were omitted from the study because it was felt that they would not possess adequate skills to read the experimental passage.

Materials

The maze tests were based on two passages of approximately 300 words each. One of these was selected from a 3¹ basal reader, and its readability was further verified by means of two formulae. The Harris-Jacobson level was "low third" and the Fry level was third grade. The second passage was taken from the sixth-grade basal reader of the same series. Again the formulae were employed, and both the Harris-Jacobson and Fry estimates were sixth grade.

For each passage, maze items were constructed for 48 of the words, regularly spaced at five-word intervals. Each item required a choice between the correct anwer and a single distractor, classifiable relative to the answer as:

1. graphically similar,same form class (GS-SFC),
2. graphically similar, different form class (GS-DFC),
3. graphically different, same form class (GD-SFC),
4. graphically different, different form class (CD-DFC).

A number of objective criteria were used to classify the distractors chosen for the tests. These guidelines were minimal and were frequently exceeded. Distractors of the same form class were of the same part of speech as the answer and agreed in number, tense, and case where appropriate. Graphically similar distractors began with the same two letters as the answer and varied by no more than three letters from the length of the answer. Graphically different distractors began with two letters different from those of the answer and varied by at least four letters from the length of the answer.

For every item, one distractor of each of the four types was chosen. The order of the distractor types was rotated in four ways producing four counterbalanced forms of the third-grade test and four of the sixth-grade test. Each test contained 12 items of each distractor type, and for any given item the type of distractor differed among the four third-grade tests and the four sixth-grade tests. A random pattern was generated for placing the distractor above or below the answer, and the same pattern was used for all eight maze tests.

Procedure

In the same week during which the *California Achievement Test* was administered, each student received one of the four alternate-form maze tests at the appropriate grade level. The printed instructions for both third and sixth graders were the same: "Whenever two words are written on top of one another, circle the word that makes sense in the sentence." In addition, the third graders were shown one example on the blackboard. It was thought best to minimize advance discussion in order to preclude any instructional effect. The testing was untimed.

Data Source

The number of correct responses made by a student served as the dependent measure. There were two levels (third grade and sixth grade) of the independent variable age. There were four levels (GS-SFC, GS-DFC, GD-SFC and GD-DFC) of the independent variable distractor type. Each student responded to twelve items at each of the four levels of distractor type.

RESULTS AND CONCLUSIONS

A 2 x 4 repeated measures ANOVA was calculated to analyze the response data. A significant main effect of age ($F\,(1{,}38) = 6.53$, $p < .05$) was observed. There was no significant main effect of distractor type ($F\,(3{,}114) = 0.31$, $p > .05$). The age x distractor type interaction observed was significant ($F\,(3{,}114) = 3.39$, $p < .05$). Post hoc comparisons showed that the significant F ratio for the interaction effect was due to differences existing between the mean scores of third-grade students on GD-SFC and GD-DFC items and the mean scores of third-grade students on GS-SFC and GS-DFC items as well as sixth-grade students on items of all four distractor types. The means and standard deviations are presented in Table 1.

TABLE 1
Means and standard deviations for correct responses as a dependent measure function of age and distractor types

	Distractor Types			
Age	**GS-SFC**	**GS-DFC**	**GD-SFC**	**GD-DFC**
Third grade	11.65	11.55	11.15	11.35
	(0.65)	(0.97)	(0.91)	(0.85)
Sixth grade	11.70	11.70	11.95	11.90
	(0.64)	(0.56)	(0.22)	(0.30)

In terms of the initial hypotheses the results indicate that sixth-grade students in the present study were significantly more accurate in their responses to maze items than were third grade students. Second, the accuracy level of responses made by all subjects in the present study was not significantly affected by distractors varying in graphic similarity or form class. Finally, the accuracy level of sixth-grade subjects to items of all four distractor types (GS-SFC = 11.7, GS-DFC = 11.7, GD-SFC = 11.95, GD-DFC = 11.9) and of third grade students to GS items (GS-SFC = 11.65, GS-DFC = 11.55) were all similar. However, the accuracy level of third-grade students to GD items (GD-SFC = 11.15, GD-DFC = 11.35) differed from the others. Thus GS items were responded to at equivalent levels of accuracy by third-and sixth-grade subjects, but third-grade subjects were not able to maintain the high level of accuracy displayed by sixth-grade subjects on GD items.

IMPLICATIONS

One general implication for maze testing must be drawn based upon the performance of students in this sample. Simply stated, maze items containing one distractor and one correct answer do not pose a challenging task to third- and sixth-grade students when they are working on materials at or slightly below their instructional reading level. Although *statistically* significant differences were discovered between third- and sixth-grade students in the present study, as well as a statistically significant interaction between age and distractor type, the educational significance of these differences is negligible. Third graders were able to respond accurately to approximately 95% of the items encountered while sixth graders were able to respond to 98% of the items accurately. The significant main effect of age and the significant interactive effect between age and distractor type observed in this study were a function of the extremely small variances, caused by uniformly high accuracy of responses across students, rather than by any sizable differences between the means. The ease with which the students completed the tasks also may have prevented any potential variance caused by the type of distractor from surfacing.

The ease of the task is more apparent if expected levels of accuracy for cloze performance are compared with the actual achieved levels of accuracy for maze performance. Considering that the students in this study were working with materials assumed to be at or slightly below their instructional level, scores of approximately 50% correct or higher would be anticipated if the passages were constructed in a cloze format. This differs markedly from the actual overall level of accuracy (96%) achieved when the passages were constructed in a maze format. Prior to this study, it was anticipated that maze items would be less difficult to complete than cloze items but the degree of the diminishment of difficulty was not anticipated. On the other hand, the fact that the maze task is one of recognition, while the cloze task is one of recall, suggests that the high maze scores recorded in this study may be analogous to the word recognition scores used as criteria in oral reading. Future studies may need to incorporate more distractors per item and require finer distinctions to be made between the correct response and the distractors if developmental trends and distractor types are to be explored as sources of variance in maze test performance of students reading materials at or below their instructional levels.

REFERENCES

Bormuth, J. R. Experimental applications of cloze tests. In J. A. Figurel (Ed.), *Improvement of reading through classroom practice*, Proceedings of the International Reading Association, 9. Newark, DE: International Reading Association, 1964.

Bradley, J., Ackerson, G., & Ames, W. The reliability of the maze procedure for classroom assessment. *Journal of Reading Behavior*, 1978, *10*, 291-296.

Bradley, J., & Meredith, K. The reliability of the maze procedure for intermediate and junior high school students. In P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice*. Twenty-Seventh Yearbook of the National Reading Conference. Clemson, SC: National Reading Conference, 1978.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. *Journal of Educational Psychology*, 1973, *65*, 294-300.

Wiseman, D., & McKenna, M. C. Classroom uses of the maze procedure. In *Selected articles on the teaching of reading*, Set C. No. 50. New York: Barnell Loft, 1978.

ROBERT B. RUDDELL
University of California, Berkeley

LITERACY ACHIEVEMENT PROFILES AND LITERACY USE — HIGH AND LOW ACHIEVERS

BACKGROUND AND RESEARCH PROBLEM

The purpose of this study is to examine the relationship between literacy achievement performance and literacy use of high and low achieving students at the junior high and senior high school levels. For the study, literacy is defined as the possession of the skills needed to perform a reading-language task imposed by self or by an external agent. This definition broadens Sticht's (1975) definition of literacy to include self-imposed literacy goals.

Given the large number of students graduating from high schools with literacy problems, with estimates ranging from 5% (Fisher, 1978) to 30% (Bormuth, 1977), it becomes clear that a literacy problem of substantial proportions exists in our education system. Increased attention must be given to improvement of literacy excellence at all levels of schooling (Sharon, 1972; Vogt, 1972).

A major conflict is present, however, in our attempt to remedy the literacy problem. While we know that students are experiencing literacy difficulties at various levels of schooling, we have very limited information as to the nature of achievement profiles and uses of literacy during the student's educational experience. The work of Sticht (1975) details literacy demands for occupations in the military, with logical extension to civilian occupations. Cook (1977) chronicles the development of adult literacy education in the United States. Evidence, however, on literacy performance and literacy use of the junior high and secondary student is virtually nonexistent.

RESEARCH QUESTIONS

The following questions served to guide the direction of the study:

1. What is the nature of reading, writing, and spelling achievement performance for high achieving and low achieving students in grades eight, nine, and ten?
2. What is the nature of in-school and out-of-school literacy used by high achieving and low achieving students?
3. What is the nature of self-perceived literacy needs and occupational goals for high achieving and low achieving students?
4. What is the nature of teacher influence on literacy development of high achieving and low achieving students based on student self-perception?

METHODOLOGY

The study was conducted in a Northern California school district in the spring of 1978. The Ss selected at grades eight, nine, and ten were part of an original sample of 522 students who had participated in an extensive language study in the 1970-71 academic year (Ruddell & Williams, 1972). Of the original sample, 193 remained in the school district — approximately 37% of the original group. Parental permission was obtained for 132 students to be excused from English classes for two testing sessions. The ethnic and SES composition of subjects across the three grade levels when compared with the 1970-71 composition revealed close similarity. The greatest difference was 10% occurring between grades 2 and 3 at semi-skilled SES level 4.

Data collected on each student at grades eight, nine, and ten consisted of the following: reading total score for the *Comprehensive Test of Basic Skills* (CTBS); vocabulary subtest for the CTBS; comprehension subtest for the CTBS; written language and spelling achievement based on the *Multiple-Picture Stimulus* from Project DELTA and evaluated by the *DELTA Written Language Inventory* which accounts for quality of thought, organization of expression, quality and control of language, fluency of language, and technical skills of written expression (Ruddell & Williams, 1972); and the *DELTA II Literacy Questionnaire* designed to provide data on in-school and out-of-school literacy use, including magazine and newspaper reading and television viewing habits, self-perceived literacy needs, occupational goals, and teacher influence on literacy development.

Inter-rater reliability for the *DELTA Written Language Inventory* was established at .90. The rating instrument had been previously used in an extensive language study (Ruddell & Williams, 1972) and was shown to be a more sensitive measure of language development than the traditionally-used measure of communication unit length.

Raw scores for the standardized achievement variables were converted to percentile scores for those high achieving students falling into approximately the top 33% and the low achieving students falling into the bottom 33% of the sample at each grade level. Profiles were then constructed on the basis of data related to reading, writing, and spelling achievement variables.

Because of the nature of the sample the Mann-Whitney U (Wilcoxon Two-Sample Test) test was used to test for significance between high and low achiever performance for the reading and writing variables. This test assumes that scores are discrete, does not assume that variances are equal, or that distribution of scores is normal (Marascuilo & McSweeney, 1977).

Data related to the *DELTA II Literacy Questionnaire* were tabulated and converted to percentages for description reporting purposes.

FINDINGS, CONCLUSIONS, AND DISCUSSION

The first research question inquired about the nature of achievement performance for high achieving and low achieving students across grades eight, nine, and ten. The CTBS vocabulary achievement for high achievers is at the 88 percentile while low achievers is at the 44 percentile. CTBS comprehension achievement for high achievers is at the 85 percentile and at the 43 percentile for low achievers. The CTBS total reading for high achievers is at the 88 percentile and at the 48 percentile for low achievers ($N = 41$). These differences were found to be significant at the .05 level ($Z = 4.01$; variance 1427.44; $U_1 = 356$; $U_2 = 52$.)

These data lend support to previous research which has established that vocabulary knowledge and ability to derive information from context using syntactic and semantic strategies are both critical to the comprehension process (Arlin, 1976; Guthrie, 1973; Leigold, 1974; Richeck, 1976; Neuwirth, 1976; Clark & Clark, 1977).

Also tied directly to the achievement profile question are the data related to writing achievement. These data reveal that the high and low achiever differs significantly on all aspects of writing evaluated namely Quality of Thought, Organization of Expression, Quality and Control of Language, Fluency of Language, and Technical Skills. The Z score equals 4.65 and is significant at the .05 level (Variance = 580.39; $U_1 = 225$; $U_2 = 0$). The finding related to the Quality and Control of Language reinforces the CTBS vocabulary achievement difference in that the writing achievement variable also strongly reflects vocabulary range, effective use of vocabulary, and knowledge of concepts.

The organization of expression variable accounts for effective use of such features as introduction, development, conclusion, sequencing, relationships, transitions, and audience awareness. An understanding of the role these features play in written expression would also appear necessary for comprehension inference. Just as word meaning and vocabulary knowledge contribute to deriving meaning from context, contextual meaning in sentences is considered an important factor in inferring text meaning (Perfetti & Goldman, 1975; Thorndyke, 1976; Garrod & Stanford, 1977; Brown & Smiley, 1977).

The longitudinal research by Ruddell (1979) has shown that reading comprehension at grades two and three contributes significantly to writing achievement at grades nine (.57) and ten (.30), respectively. It would appear that this shared variance may be accounted for in part by language fluency, organization, and control of syntactical complexity. These findings reinforce those found in the work of Evanechko, Ollila, and Armstrong (1974).

Tabulation of spelling errors was based on two samples collected from each subject. The first was based on the *Multiple Picture Stimulus* writing experience collected in 1970, as part of the longitudinal study mentioned earlier. The second was based on the identical 1978 *Multiple Picture Stimulus* writing sample. The analysis identified spelling errors on the basis of error type. A phonetic error is defined as a misspelling which closely follows a one-to-one sound symbol relationship, e.g., *fish* spelled as *fech*; a visual error is defined as a misspelling which has most or all of the letters present but arranged incorrectly, e.g., a *girl* spelled as *gril*; a classification of other includes remaining misspellings.

The largest percentage of spelling errors for both high and low achievers for the 1970 writing sample is found in the phonetic error category. The low achievers use a phonetic misspelling in 22% of their total running words contrasted with 5% for the high achievers. The 1978 phonetic error rate has decreased significantly for both groups. The low group phonetic error rate is 1.2% of running words while the high group is .05% of running words.

The visual error type for 1970 shows a 5% level for the low group and a 2.5% level for the high group. The 1978 visual error type decreases to a 1.2% level for the low group and .05% for the high group. Other category for the 1970 low group is 7.6% of running words in contrast to 1.4% for the high group. In 1978 the low group decreased to 1.4% and the high group to .10%

of running words.

The mean error rate per ten words for 1970 is 3.0 for the low group and 1.4 for the high group. Thus, the low group makes approximately twice as many errors in 1970 as does the high group per ten words. The 1978 data show the mean error per ten words to be .40 for the low group and .05 for the high group. Thus, while the overall error rate decreases for both groups over the eight-year period, the low group in grades eight, nine, and ten makes eight times as many errors per ten words as does the high group.

The spelling error rate for the low achieving group in contrast to the high achieving group lends support to the writing achievement profile data which show a significant difference in technical skills of writing, which includes spelling. It is hypotnesized that the spelling difficulties encountered by the low achieving group may directly affect the writing fluency of this group. The mental processing energy required to concentrate on correct spelling may directly affect fluency, organization, and quality of written expression. The five language factors would appear to be highly correlated.

Question two of the study inquired about the nature of in-school and out-of-school literacy use by high achieving and low achieving students. Several parts of the questionnaire dealt with this question, including literacy use related to academic interests, personal interests, magazine reading, newspaper reading, television viewing, and work experience.

Academic related literacy use — defined as use of reading, writing, or speaking skills in school or in completing homework — revealed that high achievers use reading skills more frequently than low achievers in reading directions (75% vs 61%) (Percent of high achieving subjects will be reported first, followed by low achieving subjects unless otherwise indicated.) and taking tests (57% vs 44%); use writing skills more frequently in taking tests (61% vs 50%); use speaking skills more frequently in asking questions (82% vs 67%) and in club meetings (7% vs 0%). On the other hand, low achievers use reading skills more frequently than high achievers in reading textbooks (94% vs 86%), library resource books (50% vs 36%), and free-reading books (39% vs 18%); use writing skills more frequently in homework assignments (100% vs 75%) and in completing papers and reports (67% vs 54%); and use speaking skills more frequently in group discussions (56% vs 46%). Both groups used the following literacy skills in near identical fashion: reading class notes (72% vs 71%); writing class notes (78% vs 77%); writing stories (39% vs 33%) and poetry (7% vs 5%); speaking to answer questions (64% vs 61%); oral reports (39% vs 39%); and telling stories (18% vs 22%).

Seventy-one percent of the high achievers report that they spend one hour or less on homework in contrast to 55% of the low achievers. On the other hand, 45% of the low achievers spend two or more hours in after-school study in contrast to 29% of the high achievers.

While the problem inherent in self-reporting must be recognized, it would appear that high achievers make greater use of literacy skills in academic areas with high performance visibility such as test taking, direction reading, question asking, and in club meetings. Low achievers appear to spend more time using literacy skills in reading text and library resource books, on homework, and in completing papers and reports. These differences follow a logical pattern suggesting that the high achiever may be more efficient in using literacy skills in academic pursuits during the school day while the low achiever requires extended literacy skill use in completing text, resource reading, and homework assignments.

The question which probed the out-of-school personal interest reading for the students indicated that high achievers used reading more than low achievers in reading magazines (75% vs 67%) and recipes (36% vs 28%). Writing use was used predominately more by high achievers in writing directions (28% vs 17%), lists (54% vs 44%), stories (64% vs 39%), and poetry (43% vs 6%) while speaking skills were used more in group meetings (21% vs 11%) and in telling stories (32% vs 22%). A high degree of similarity was found between the two groups in reading notes (28% vs 22%), directions (46% vs 44%), resource information (57% vs 50%), and books (64% vs 61%). Similar use was reported by both groups in letter writing (68% vs 67%). Almost identical use was reported by both groups in family conversation (82% vs 88%), asking questions (71% vs 72%), and in answering questions (68% vs 72%).

These differences suggest that while similarity is present in using literacy skills in out-of-school personal interests, the higher achievers use literacy skills substantially more in magazine and recipe reading, writing more in formulating directions, lists, stories, and poetry, and speaking more in group meetings and telling stories. These differences may reflect greater literacy skill transfer for the higher achiever but may also reflect a socioeconomic difference and extended opportunities for literacy use in the home environment.

In response to a question designed to identify magazine reading either in or out-of-school, at least 25% of the low achievers identified 11 of 49 popular magazine selections. The data indicate that the low achievers read these magazines at a substantially higher frequency than do high achievers. These consisted of *Newsweek* (39% vs 11%), *Time* (33% vs 18%), *Sports Illustrated* (56% vs 18%), *Road and Track* (28% vs 7%), *Better Homes and Gardens* (33% vs 7%), *TV Guide* (33% vs 21%), *Reader's Digest* (28% vs 7%), *Ebony* (61% vs 11%), *Mad* (44% vs 32%), *People* (44% vs 21%), and movie magazines (28% vs 7%). The high achieving group report reading only one magazine at a substantially higher degree than the low achieving group — *National Geographic* (64% vs 6%). It would appear that the popular magazines cited hold much greater appeal for the lower achiever than the high achiever. Many of these periodicals are liberally illustrated and are written at a comparatively low readability level. The pattern of reading interest may thus follow an expected literacy use pattern related to achievement level.

In response to the question, "Do you read a daily newspaper?" 71% of the high achieving group responded positively as did 83% of the low achieving group. Approximately equal time was spent in reading front page news stories (71% vs 72%) and classified ads (29% vs 28%). High achieving subjects more frequently read the comics (75% vs 61%) and by-lined columns (39% vs 17%) than did low achieving subjects. The low achieving subjects read the sports section (61% vs 46%) and the religious section (11% vs 0%) to a greater extent. Personal interests would seem to influence this selection process to a greater extent than academic achievement.

The last questionnaire item which related to out-of-school literacy use dealt with part-time job activities. In the high group 56% reported part-time job activities while 44% of the low group worked after school or on weekends. Jobs for both groups range from batboy for the CAL baseball team and washing dishes to babysitting and peeling apples at Fat Albert's, a local pie shop.

Sixty percent of the high achieving subjects reported they spent approximately 15 minutes or less on a daily basis in using literacy skills on their job in contrast with 26% of the low achievers. Forty percent of the high achievers indicated 30 to 60 minutes time utilized on the job in contrast to 74% of the low achievers.

Part-time job experience involving the use of reading, writing, or speaking skills indicated that the high achievers use reading skills more frequently than low achievers in reading directions (65% vs 38%), orders (73% vs 50%), information (40% vs 13%) and directions (67% vs 38%); and speaking in answering telephones (53% vs 38%); asking questions (60% vs 50%), telephoning customers (33% vs 0%), giving directions (47% vs 38%), telling stories (20% vs 13%), and making presentations (20% vs 0%). Low achievers use literacy skills more frequently in reading notes (38% vs 27%); and writing lists (38% vs 27%) and orders (25% vs 13%). Approximate similarity was present using speaking skills for both groups in answering questions (60% vs 63%) and talking to customers (47% vs 50%).

These findings suggest that the high achievers more frequently apply their literacy skills to the part-time job setting than do low achievers. The high achievers report using literacy skills during a shorter time interval than do low achievers; however, the nature of the tasks reported suggests that the higher achievers may simply be more efficient in application of skills. Assuming this hypothesis is correct, it parallels the use efficiency of the high achiever previously observed in the academic use of literacy skills.

The third question of the study inquired about the nature of self-perceived literacy needs and occupational goals for high and low achieving students. High achievers more frequently expressed the view that teachers should place greater emphasis on reading for pleasure (7% vs 0%); story writing (89% vs 36%); presenting oral reports (75% vs 50%); and speaking with adults (61% vs 44%) than low achievers. Low achieving Ss more frequently indicated that emphasis should be placed on reading, specifically vocabulary study (83% vs 68%) and reading comprehension (78% vs 61%); on writing with attention to spelling study (78% vs 46%) and handwriting drill (56% vs 25%); and on speaking emphasizing asking good questions (83% vs 68%). Both groups similarly identified reading emphasis on phonics (36% vs 28%) and oral reading (71% vs 78%); writing stressing report writing (68% vs 72%) and punctuation study (39% vs 44%); and speaking with emphasis on participation in discussions (78% vs 71%).

These self-identified needs logically reflect the achievement profiles previously discussed. The low achiever expresses specific need for increased emphasis on vocabulary study, reading comprehension, asking good questions, report writing, spelling, and handwriting. The high achiever expresses greater interest in oral reporting, pleasure reading, speaking with adults, and story and report writing. Both groups express high need for greater emphasis on participating in discussion.

Occupational goals identified by both groups of students ranged from veterinarian and surgeon to police officer and wine-maker. Using the *California Socio-Economic Scale of Urban Occupations* developed by Ruddell and Williams (1972), an occupational mean was obtained for each group.

The scale ranks occupations across a five-point scale. A rank of one represents the highly-trained professional, unique individuals of national prominence and owners of major establishments; a rank of two represents less highly-trained professionals including teachers, librarians, researchers, nurses, and computer programmers; through a rank of three and four, representing skilled and semi-skilled workers respectively, to a rank of five, representing unskilled workers such as day-laborers, janitors, dishwashers, nurses' aides, and welfare recipients. The mean ranking of the high achievers was 1.8 and contrasts only slightly with the low achievers with a rank of 2.1. Occupational aspirations thus vary little for the two groups. This finding parallels that of previous research (Bloom, 1976) which indicates that aspirations for the low achiever may be unrealistically high.

The fourth and final research question explored the nature of teacher influence on literacy development of the high and low achieving students, based on student self-perception. All of the high achieving students identified at least one teacher that had influenced them while only 72% of the low achievers provided a response. Based on the analysis of the responses, five teacher influence categories were formulated. High and low achiever response to the categories is remarkably similar. Personal characteristics of the teacher were identified by 14% of high achiever and 13% of low achiever responses. Teacher understands learner potential received 7% of high achiever and 4% of low achiever responses. Teacher helped develop positive attitude toward subject was identified by 15% of high and 13% of low achiever responses. The category helped make life adjustment was identified by 19% of the high achievers and 23% of the low achievers. Quality of instruction was identified by 45% of high achievers and 47% of low achievers. For high achievers 17% of the teachers were found in grades kindergarten through six and 83% in grades seven through ten, contrasted with 13% of grades kindergarten through six and 87% at grades seven through ten for the low achievers.

Even though high and low achiever response appears to be similar, a more significant factor in the student response pattern may be found in the response frequency. The high achievers provided an average of 3.2 responses each while the low achievers provided an average of 1.7 per student. Thus, while high and low achiever responses are distributed in a similar manner across the five teacher categories, the high achievers identified almost twice as many teachers who were perceived as having great influence on their academic development. This finding may reflect the high value priority which teachers place on high achievement which may, in turn, reinforce the high achieving student and teacher's concept of self-worth in the instructional setting.

SUMMARY — RESEARCH RECOMMENDATIONS

In summary several general conclusions may be identified. First, the achievement profiles for high achievers differ significantly from low achievers in reading vocabulary, reading comprehension, and total reading. Distinct differences are found in writing expression including quality of thought, organization of expression, quality and control of language, fluency of language and technical skills such as spelling. These findings support the idea that the high and low achievement variation may be attributed to such factors as vocabulary knowledge, sentence inference, and text organization or story schemata, all of which affect the "unfolding" of information contained in the text.

But while the profiles of high and low achievers vary in achievement level, they are remarkably similar in relative distribution. The reading and writing variables would appear to be highly correlated even though progress of the low achievers is at a slower pace. One must question how instruction at the junior high and secondary school level can best assist the low achiever if reading and writing skills are to be improved over time. How can the teacher at these levels best provide for differential achievement levels? As Bloom has pointed out "if students are normally distributed with respect to *aptitude* for some subject and all students are given exactly the *same* instruction (the same in terms of amount and quality of instruction and learning time allowed), then achievement measured at the completion of the subject will be normally distributed. . . . Conversely, if students are normally distributed with respect to aptitude, but the kind and quality of instruction and learning time allowed are made appropriate to the characteristics and needs of *each* learner, the majority of students will achieve mastery of the subject. And, the correlation between aptitude measured at the beginning of instruction and achievement measured at the end of instruction should approach zero." (1976, p. 4)

Second, literacy use of high and low achievers differs substantially for both in-school and out-of-school activities. These differences suggest the high achievers exhibit a high level of literacy efficiency, in both academic and work settings. The high achiever more frequently selects magazines with high vocabulary loadings, than is the case with the low achiever.

Third, self-perceived literacy needs of both high and low achievers mirror academic performance. The high achiever perceives greater need for instructional emphasis on story writing, reading for pleasure, presenting oral reports and speaking with adults; while the low achiever expresses greater need for emphasis on vocabulary study, reading comprehension, spelling, handwriting, and asking good questions. Occupational goals are similar for both high and low achievers and may represent unrealistic expectations for the low achiever.

Fourth, high and low achievers perceived teachers who had influenced their academic and personal development in a remarkably similar way. The major difference however, resided in the frequency of response. High achievers identified almost twice as many significant influence teachers as did the low achievers. Teachers appear to express a significantly greater degree of interest in the high achiever than in the low achiever which may in turn affect achievement motivation for the high achiever.

This descriptive research has been designed to provide information leading to a better understanding of the relationship between literacy achievement and literacy use in both in-school and out-of-school environments. Future research must probe deeper into this relationship through intensive classroom and work observations, and follow-up interviews with both students and teachers. The appropriateness and quality of instruction and the variable of learning time must be carefully researched if we are to better understand the most effective means of developing high levels of literacy with both low and high achieving students.

ACKNOWLEDGEMENT

The author acknowledges Mr. Lee Gunderson, Graduate Assistant, University of California, for his support in data tabulation and reaction.

REFERENCES

Arline, M. Causal priorities between comprehension subskills: word meaning and paragraph meaning. *Journal of Reading Behavior*, 1976, *8*, 351-361.

Bloom, B. S. *Human characteristics and school learning*. New York: McGraw-Hill, 1976.

Bormuth, J. R. Literacy policy and reading and writing instruction. In R. Beach & P. D. Pearson (Eds.), *Perspectives on literacy*., Minneapolis, Minnesota: University of Minnesota, 1977.

Brown, A. L., & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 1977, *8*, 1-8.

Clark, H. H., & Clark, E. V. *Psychology and language*. New York: Harcourt Brace Jovanovich, Inc., 1977.

Cook, W. D. Adult literacy education in the United States. Newark, Delaware: International Reading Association, 1977.

Evanechko, P., Ollila, L., & Armstrong, R. An investigation of the relationships between children's performance in written language and their reading ability. *Research in the Teaching of English*, 1974, *8* (3), 315-326.

Fisher, D. L. *Functional literacy and the schools*. Washington, D.C.: The National Institute of Education, 1978.

Garrod, S., & Stanford, A. Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 77-90.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. *Journal of Educational Psychology*, 1973, *66*, 294-299.

Lesgold, A. M. Variability in children's comprehension of syntactic structures. *Journal of Educational Psychology*, 1974, *66*, 333-338.

Marascuilo, L., & McSweeney, M. *Nonparametric and distribution-free methods for the social sciences*. Monterey, California: Brooks/Cole Publishing Company, 1977.

Neuwirth, S. E. A look at intersentence grammar. *Reading Teacher*, 1976, *30*, 28-32.

Perfetti, C. A., & Goldman, S. R. Discourse functions of thematization and topicalization. *Journal of Psycholinguistic Research*, 1975, *4*, 257-271.

Richek, M. A. Effect of sentence complexity on the reading comprehension of syntactic structures. *Journal of Educational Psychology*, 1976, *68*, 800-806.

Ruddell, R. B. Early prediction of reading success: Profiles of good and poor readers. In M. L. Kamil & A. J. Moe (Eds.), *Reading research: Studies and applications*. Twenty-eighth Yearbook of the National Reading Conference. Clemson, S.C.: National Reading Conference, 1979.

Ruddell, R. B., & Williams, A. C. *A research investigation of a literacy teaching model: Project DELTA* (EPDA Project No. 005262). Washington, D.C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1972.

Sharon, A. T. *Reading activities of American adults*. Princeton, New Jersey: Educational Testing Service, 1972.

Sticht, T. G. (Ed.). *Reading for working: A functional literacy anthology*. Alexandria, Virginia: Human Resources Research Organization, 1975.

Thorndyke, P. W. The role of inference in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 437-446.

Vogt, D. *Literacy among youths 12-17*. Vital and Health Statistics Series 11-131. Washington, D.C.: U.S. Government Printing Office, 1973.

TIMOTHY SHANAHAN
University of Delaware

THREE EXPLANATIONS OF READING DISABILITY

The purpose of this paper is to examine three theoretical explanations of reading disability from the vantage of an objectively collated collection of published empirical studies which have used a good reader-poor reader comparison paradigm. The ability of visual, linguistic, and developmental theories of reading disability to explain or predict phenomena described by extant data will be tested. Those specific processing characteristics are emphasized which have been suggested as discriminating factors that distinguish good and poor readers. The intention is to explore the phenomena, and not the etiology of reading disability.

METHOD

(1) The Current Index to Journals of Education (CIJE) was accessed for all research articles concerning Reading Difficulty, Dyslexia, Reading Processes, and Reading Acquisition published between 1970-78. (2) The citations from CIJE were examined and only those studies dealing with cognitive, perceptual, or linguistic aspects of reading disability which uses good reader-poor reader comparisons were examined. (3) The bibliographies of the CIJE articles were searched for additional relevant citations. In all, 74 studies were reviewed for this paper. Although this body of research is not comprehensive, it is both objectively compiled and extensive.

AN ANALYSIS OF THE VISUAL-PERCEPTUAL EXPLANATION OF READING DISABILITY

The earliest investigations of reading disability were produced by physicians, often opthalmologists. These studies usually focussed upon the visual or perceptual aspects of the reading process. This visual-perceptual approach to the study of reading disability is still popular. Recent studies of the visual-perceptual characteristics of poor readers have examined icon persistence, perceptual span, visual encoding speed, visual memory, perceptual accuracy, and auditory-visual integration.

Icon Persistence

Research indicates that the visual image or icon created by the visual process must be retained for an adequate length of time to be processed (Sperling, 1960; Kahneman, 1966). If additional visual information intrudes too rapidly, the initial icon cannot be analyzed. Although studies (Riding & Pugh, 1977; Stanley & Hall, 1973) have reported significant differences in the separation times required to perceive visual images for good and poor readers (i.e. poor readers require more separation between inputs), it is improbable that such differences alone would be meaningful. Under ordinary circumstances, the eyes would adust to this by making longer fixations, and thus, the poor readers would get the same visual information as good readers, though they would read more slowly. Unless this slower reading was found to produce a memory overload, it probably is not an important factor in reading disability.

Perceptual Span

Another aspect of visual processing which might be related to reading disorders concerns the perceptual span, or the amount of information which can be collected in a single fixation. Studies which have attempted to examine the perceptual spans of good and poor readers have indicated the existence of significant differences (Spring & Farmer, 1975; Bouma & Legein, 1975), although no controls for IQ differences were used. The differences found in these studies were apparent only with children under the age of ten, and were so limited (i.e. about one letter space difference per fixation) that it is doubtful that they provide an adequate explanation of reading disability.

Visual Encoding Speed

If the visual image is transmitted so slowly to the brain that mnemonic processing is disrupted, visual encoding speed might provide an explanation of reading disability. Studies of encoding speed have produced mixed results, however (Lyle & Goyen, 1975; Morrison, Giordani, & Nagy, 1977). The inconsistent results may have been due to (1) differences in the ages of subjects studied, and (2) a fatigue factor. Spring (1971) reports that visual processing differences in good and poor readers do not become apparent until after about ten minutes of testing has been completed; poor readers are more apt to be inattentive to reading type tasks after that time limit.

Visual Memory

The ability to retain a visual percept long enough to analyze it in relationship to other information is an important requisite skill in reading. Poor readers have been found to be somewhat deficient in visual memory, but usually only with strings of five or six letters. Most attempts to examine visual memory have been confounded by a lack of control for differences in verbal rehearsal strategies. Although the memory differences appear to be small, they might be particularly important with younger children (Cummings & Faw, 1976; Lyle & Goyen, 1968; Goyen & Lyle, 1973; Vellutino, Smith, Steger, & Kaman, 1978; Lahey & Lefton, 1977; Allington, Gormley, & Truex, 1976; Vellutino, Steger, Kaman, DeSetto, 1975). Future studies of visual memory differences need to attempt to separate the influence of verbal mnemonic strategies from visual memory.

Perceptual Accuracy

It has often been proposed that poor readers misperceive visual information; clinical reports often indicate that poor readers see reverse images of words and letters. However, poor readers were found in several studies to be able to copy words better than they had read them, an impossible task if the misreadings were due to misperception; poor readers were found to make accurate choices in visual search tasks; and, only a small percentage of the oral reading errors of poor readers were of the reversal variety (Vellutino, Smith, Steger, & Kaman, 1975; Lahey & Lefton, 1977; Allington, Gormley, & Truex, 1976; Clifton-Everest, 1975; Lyle, 1969; Liberman, Shankweiler, Orlando, Harris, & Berti, 1971). Poor readers do not appear to be deficient in perceptual accuracy.

Auditory-Visual Integration (AVI)

Possibly the most thoroughly researched aspect of visual-perceptual deficit is the question of auditory-visual integration. A deficiency in the ability to match incoming auditory and visual information would not disrupt the reading process, but it could seriously interfere with learning to read. Available data indicates, however, that AVI problems probably have little to do with reading disability (Birch & Belmont, 1965; Bryden, 1972; VandeVoort, Senf, & Benton, 1972; Rodenborn, 1970-71; Blank, Weider, & Bridger, 1968; Kastner & Rickards, 1974; Vellutino, Steger, & Pruzek, 1973; Vellutino, Pruzek, Steger, & Meshoulam, 1972; Vellutino, Harding, Phillips, & Steger, 1975; VandeVoort & Senf, 1973; Badian, 1977). Deficiencies in AVI task performance are probably due to variations in verbal memory, attention, or rehearsal strategies, and not to some neural confusion.

AN ANALYSIS OF THE VERBAL-DEFICIT EXPLANATION OF READING DISABILITY

A more recently proposed explanation of reading disability is that it is due, not to visual-perceptual deficiencies, but to a limitation in the ability to store and retrieve linguistic referents from verbal memory. Studies of verbal processing differences have considered the importance of verbal intelligence, as well as more specific processing deficits in vocabulary knowledge, verbal memory and rehearsal, and syntax.

Verbal Intelligence

Reading is, at least in part, a language process. Deficiencies in the ability to process language could have a disruptive influence on reading development. Studies indicate that poor readers score significantly lower on tests of verbal intelligence than do good readers (Belmont & Birch, 1965). Even when good and poor readers are matched in terms of full scale intelligence scores, the proficient readers score significantly higher than the poor readers on the verbal scale. There is also a greater discrepancy between the performance and verbal intelligence scores obtained by poor readers, though this is probably just artifact. Why would a child with average verbal IQ scores and above average performance scores have difficulty learning to read? These verbal IQ differences do suggest that poor readers may have difficulty because of verbal processing problems. These differences also indicate that paradigms which match good and poor readers for verbal IQ scores probably serve to attenuate important differences which separate these groups in a random sample.

Vocabulary

Studies of differences in word meaning knowledge have attempted to provide an explanation of why verbal intelligence might be deficient in poor readers. Several studies have indicated that the clearest IQ test subscore divergence between good and poor readers was usually found on the vocabulary subtest (Belmont & Birch, 1965; Lyle & Goyen, 1969; Bell & Aftanas, 1972; Levine & Fuller, 1972; Muehle & DiNello, 1976). It has also been demonstrated that the discrepancies between the vocabulary performance of good and poor readers do not increase from year to year as a result of reading instruction (Lyle & Goyen, 1969), indicating that reading disability probably does not cause vocabulary knowledge to be deficient. Golinkoff (1974-75), from a literature review, indicates that a vocabulary deficit may explain some aspects of reading disability, but it probably fails to account for more important aspects of reading problems.

Verbal Memory

Although vocabulary test scores are relatively low for disabled readers, this is not the only IQ subtest that reflects a difference. In fact all subtests, including vocabulary, which are largely dependent upon memory, separate good and poor readers (Torgeson, 1978-79). Poor readers have been found to be slower in all rapid serial recall or naming tasks which require lexical access from memory (Denckla, 1972; Denckla & Rudel, 1976; Eakin & Douglas, 1971; Spring & Capps, 1974; Mackworth & Mackworth, 1974). This lexical access, or recall, is limited in poor readers even when no verbal intelligence differences separate the groups.

Poor readers are not only slower in accessing lexical items from memory, but they are also less adept in the use of verbal rehearsal strategies (Kastner & Rickards, 1974; Torgeson & Goldman, 1977). Whether poor readers can be taught to employ rehearsal strategies which would increase their verbal memory is unknown, though no benefits were found in one study which attempted to provide categorical organization for poor readers (Parker, Freston, & Drew, 1975). Studies which actually train the subjects to use categorical organization, or verbal rehearsal, might be more successful.

Two studies suggest that poor readers are less attentive to visual information than are good readers (Guthrie & Tyler, 1976; Kolers, 1975). No studies were found which attempted to analyze the relationship of memory and attention in poor readers. It is difficult to separate attention and memory strategies, however, as the use of rehearsal may improve reading by simply preventing the student from withdrawing attention prematurely.

Syntax

Only one other aspect of language processing, use of syntax, was examined in any of these studies. Denner (1976) found that in an artificial language processing task, poor readers had difficulty in combining sets of symbols into meaningful prose; but this may have simply been the results of the memory demands of the task, and no intelligence matching was done. A second study which examined the influence of syntactic and semantic structures on recall found poor readers to benefit from these structures as much as good readers (Guthrie & Tyler, 1976). However, this study used a poor measure of comprehension, verbatim sentence recall. It would be premature to draw conclusions about the syntactic processing abilities of poor readers.

AN ANALYSIS OF A DEVELOPMENTAL EXPLANATION OF READING DISABILITY

The developmental theory provides the most explicit explanation of reading disorders of the approaches discussed here. Those who interpret reading problems developmentally claim that there are definite, qualitatively different, stages of poor reading. The earliest stage is hypothesized to be closely related to visual-perceptual problems, and the later stage is more cognitive-linguistic in nature. The developmental qualities of reading disability are interpreted as being the result of a maturational lag in the learner (Satz & Sparrow, 1970) or as the direct result of the specific task demands inherent in learning to read (Vernon, 1977).

Only four studies were found which specifically attempted to test the hypotheses that younger children who have difficulty in reading have difficulty in perception, while older children with reading problems have verbal processing difficulties (Satz, Friel & Rudegair, 1974; Sabatino & Hayden, 1970; Satz, Radin, & Ross, 1972; Vellutino, Steger, & Kandel, 1972). These studies fail to provide much support for the developmental explanation. (1) The initial perceptual deficit does appear to decline on some measures over time. However, this appears to be the result of learning and not development. The only tasks which reflect a perceptual improvement over time are those which are measured by tests with low ceilings, such as letter learning. If a child learns the names of the letters does this indicate that the perceptual deficit which made such learning difficult has disappeared? Probably not, though a letter-naming task would probably no longer be a good measure of perceptual efficiency. (Other developmental studies also indicate that the performance of poor readers changes over time, but as the result of learning and not neutral development (Guthrie, 1973; Guthrie & Seifert, 1977)). (2) These studies were inconsistent in the attribution of perceptual problems to the younger poor readers. (3) None of these studies indicated that young poor readers do not suffer from the verbal deficits predicted to be important in the older population. The post-hoc analysis of several studies which use younger and older poor readers was undertaken to evaluate the developmental hypothesis (Muehl & DiNello, 1971; Bell & Aftanas, 1972; Lyle & Goyen, 1975; Lahey & Lefton, 1976; Allington, Gormley, & Truex, 1976; VanMeel, Vlek, & Brujel, 1970; Morrison, Giordani, & Nagy, 1977; Vellutino, Steger, DeSetto, & Phillips, 1975; Spring & Farmer, 1975; Denckla & Rudel, 1976; Kastner & Rickards, 1974; Denner, 1970). Again, this developmental explanation fails to account for differences which exist in language processing at all levels of development.

CONCLUSIONS

On the basis of the analysis of an objectively collated corpus of research, it is concluded that: (1) The most powerful, reliable explanation of reading disability is that it is caused by verbal memory deficiencies. Several studies indicated that poor readers were less apt to employ purposeful rehearsal strategies, and that they were less efficient in recalling information from memory. (2) Although visual-perceptual processes have been studied for over seventy years it appears that it would be premature to discontinue research in this area. Poor readers do not appear to misperceive visual information, but their visual processing does seem to be inefficient. Studies need to be undertaken to examine the interactions of this visual inefficiency with the verbal memory limitations of the poor reader. (3) The developmental hypothesis examined here can be rejected. Poor readers appear as if they overcome initial perceptual deficits over time, but this is probably due to the nature of the tasks and not the learners.

REFERENCES

Allington, R. L., Gormley, K., & Truex, S. Poor and normal reader's achievement on visual tasks involving high frequency, low discriminability words. *Journal of Learning Disabilities*, 1976, *9*, 292-296.

Badian, N. A. Auditory-visual integration, auditory. memory, and reading in retarded and adequate readers. *Journal of Learning Disabilities*, 1977, *10*, 108-114.

Bell, A. E. & Aftanas, M. S. Correlates of reading retardation. *Perceptual and Motor Skills*, 1972, *35*, 659-667.

Belmont, L. & Birch, H. G. The intellectual profile of retarded readers. *Perceptual and Motor Skills*, 1966, *22*, 787-816.

Birch, H. G. & Belmont, L. Auditory-visual integration, intelligence and reading ability of school children. *Perceptual and Motor Skills*, 1965, *20*, 295-305.

Blank, M., Weider, S., & Bridger, W. Verbal deficiencies in abstract thinking in early reading retardation. *American Journal of Orthopsychiatry*, 1968, *38*, 823-834.

Bouma, H. & Legein, C. P. Foveal and parafoveal recognition of letters and words by dyslexics and average readers. *Neuropsychologia*, 1977, *15*, 69-80.

Bryden, M. Auditory-visual and sequential-spatial matching in relation to reading ability. *Child Development*, 1972, *43*, 824-832.

Clifton-Everest, I. M. Dyslexia: Is there a disorder of visual perception? *Neuropsychologia*, 1976, *14*, 491-494.

Cummings, E. M. & Faw, T. Short-term memory and equivalence judgements in normal and retarded readers. *Child Development*, 1976, *47*, 286-289.

Denckla, M. B. Color-naming defects in dyslexic boys. *Cortex*, 1972, *8*, 164-176.

Denckla, M. B. & Rudel, R. G. Rapid 'automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 1976, *14*, 471-479.

Denner, B. Representational and syntactic competence of problem readers. *Child Development*, 1970, *41*, 881-887.

Eakin, S. & Douglas, V. I. "Automatization" and oral reading problems in children. *Journal of Learning Disabilities*, 1971, *4*, 26-33.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly*, 1975-76, *11*, 623-659.

Goyen, J. D. & Lyle, J. G. Short-term memory and visual discrimination in retarded readers. *Perceptual and Motor Skills*, 1973, *36*, 403-408.

Guthrie, J. T. Models of reading and reading disability. *Journal of Educational Psychology*, 1973, *65*, 9-18.

Guthrie, J. T. & Seifert, M. Letter-sound complexity in learning to identify words. *Journal of Educational Psychology*, 1977, *69*, 686-696.

Guthrie, J. T. & Tyler, S. J. Psycholinguistic processing in reading and listening among good and poor readers. *Journal of Reading Behavior*, 1976, *8*, 415-426.

Kahneman, D. Method, findings and theory in studies of visual masking. *Psychological Bulletin*, 1968, *70*, 404-425.

Kastner, S. B. & Rickards, C. Mediated memory with novel and familiar stimuli in good and poor readers. *Journal of Genetic Psychology*, 1974, *124*, 105-113.

Kolers, P. A. Pattern-analyzing disability in poor readers. *Developmental Psychology*, 1975, *11*, 282-290.

Lahey, B. B. & Lefton, L. A. Discrimination of letter combinations in good and poor readers. *Journal of Special Education*, 1976, 10, *205-210*.

Levine, M. & Fuller, G. Psychological, neuropsychological, and educational correlates of reading deficit. *Journal of Learning Disabilities*, 1972, *5*, 563-571.

Liberman, I. Y., Shankweiler, D., Orlando, C., Harris, K. S., & Berti, F. B. Letter confusion and reversal of sequence in the beginning reader: Implications for Orton's theory of developmental dyslexia. *Cortex*, 1971, *7*, 127-142.

Lyle, J. G. Reading retardation and reversal tendency: A factorial study *Child Development*, 1969, *40*, 833-843.

Lyle, J. G. & Goyen, J. Visual recognition, developmental lag, and strephosymbolia in reading retardation. *Journal of Abnormal Psychology*, 1968, *73*, 25-29.

Lyle, J. G. & Goyen, J. Performance of retarded readers on the WISC and educational tests. *Journal of Abnormal Psychology*, 1969, *74*, 105-112.

Lyle, J. G. & Goyen, J. D. Effect of speed of exposure and difficulty of discrimination on visual recognition of retarded readers. *Journal of Abnormal Psychology*, 1975, *84*, 674-676.

Mackworth, J. F. & Mackworth, N. H. How children read: Matching by sight and sound. *Journal of Reading Behavior*, 1974, *6*, 295-303.

Morrison, F. J., Giordani, B., & Nagy, J. Reading disability: An information-processing analysis. *Science*, 1977, *97*, 77-79.

Muehl, S. & DiNello, M. C. Early first-grade skills related to subsequent reading performance: A seven year followup study. *Journal of Reading Behavior*, 1976, *8*, 67-81.

Parker, T. B., Freston, C. W., & Drew, C. J. Comparison of verbal performance of normal and learning disabled children as a function of input organization. *Journal of Learning Disabilities*, 1975, *8*, 386-392.

Riding, R. J. & Pugh, J. C. Iconic memory and reading performance in nine-year-old children. *British Journal of Educational Psychology*, 1977, *8*, 132-137.

Rodenborn, L. V. The importance of memory and integration factors to oral reading ability. *Journal of Reading Behavior*, 1970-71, *3*, 51-59.

Sabatino, D. A. & Hayden, D. L. Psycho-educational study of selected behavioral variables with children failing the elementary grades. *Journal of Experimental Education*, 1970, *38*, 40-57.

Satz, P., Friel, J., & Rudegair, F. Differential changes in the acquisition of developmental skills in children who later become dyslexic. In D. G. Stein, J. R. Rosen, & N. Butter (Eds.), *Plasticity and recovery of function in the central nervous system*. New York: Academic Press, 1974.

Satz, R., Radin, D., & Ross, J. An evaluation of a theory of specific developmental dyslexia. *Child Development*, 1971, *42*, 2009-2021.

Satz, P., & Spiro, S. S. Specific developmental dyslexia: A theoretical formulation. In D. J. Bakker & P. Satz (Eds.), *Specific Reading Disability*. Netherlands: Rotterdam University Press, 1970.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, *74*, Whole No. 498.

Spring, C. Perceptual speed in poor readers. *Journal of Educational Psychology*, 1971, *62*, 492-500.

Spring, C. & Capps, C. Encoding speed, rehearsal, and probed recall of dyslexic boys. *Journal of Educational Psychology*, 1974, *66*, 780-786.

Spring, C. & Farmer, R. Perceptual span of poor readers. *Journal of Reading Behavior*, 1975, *7*, 297-305.

Stanley, G. & Hall, R. Short-term visual information processing in dyslexics. *Child Development*, 1973, *44*, 841-844.

Torgeson, J. & Goldman, T. Verbal rehearsal and short-term memory in reading disabled children. *Child Development*, 1977, *48*, 56-60.

Torgeson, J. K. Performance of reading disabled children on serial memory tasks: A selective review of recent research. *Reading Research Quarterly*, 1978-79, *14*, 57-87.

VandeVoort, L. & Senf, G. M. Audiovisual integration in retarded readers. *Journal of Learning Disabilities*, 1973, *6*, 170-179.

VandeVoort, L., Senf, G. M., & Benton, A. L. Development of audiovisual integration in normal and retarded readers. *Child Development*, 1972, *43*, 1260-1272.

VanMeel, J. M., Vlek, C. A. J., & Brujel, R. M. Some characteristics of visual information-processing in children with learning difficulties. In D. J. Bakker & P. Satz (Eds.), *Specific Reading Disability*. Netherlands: Rotterdam University Press, 1970.

Vellutino, F. R., Harding, C. J., Phillips, F., & Steger, J. A. Differential transfer in poor and normal readers. *Journal of Genetic Psychology*, 1975, *126*, 3-18.

Vellutino, F. R., Pruzek, R., Steger, J. A., Meshoulam, U. Immediate visual recall in poor and normal readers as a function of orthographic-linguistic familiarity. *Cortex*, 1973, *9*, 368-384.

Vellutino, F. R., Smith, H., Steger, J. A., & Kaman, M. Reading disability: Age differences and the perceptual deficit hypothesis. *Child Development*, 1975, *46*, 487-493.

Vellutino, F. R., Steger, J. A., DeSetto, L., & Phillips, F. Immediate and delayed recognition of visual stimuli in poor and normal readers. *Journal of Experimental Child Psychology*, 1975, *19*, 223-232.

Vellutino, F. R., Steger, J. A., Kaman, M., & DeSetto, L. Visual form perception in deficient and normal readers as a function of age and orthographic-linguistic familiarity. *Cortex*, 1975, *11*, 22-30.

Vellutino, F. R., Steger, J. A., & Kandel, G. Reading disability: An investigation of the perceptual-deficit hypothesis. *Cortex*, 1972, *8*, 106-118.

Vellutino, F. R., Steger, J. A., & Pruzek, R. M. Inter- vs. intrasensory deficit in pairing associate learning in poor and normal readers. *Canadian Journal of Behavioral Science*, 1973, *5*, 111-123.

Vernon, M. D. Varieties of deficiency in the reading process. *Harvard Educational Review*, 1977, *47*, 396-410.

VICTOR FROESE
University of Manitoba

RAUDING AND SEERING: JUDGING GLOBAL READABILITY

Estimating the difficulty of reading materials is an important teacher function. Readability procedures have varied from detailed (Dale-Chall, 1948; Spache, 1953; Bormuth, 1966), to simple (Fry, 1968; McLaughlin, 1969; Raygor, 1977; Baldwin & Kaufman, 1979), to linguistic (Granowsky & Botel, 1974), to global (Carver, 1975-76; Singer, 1975), and each procedure has claimed either reliability, validity or calculation ease as an advantage.

Because of the validity and the calculation ease reported by two global readability procedures, the Rauding Scale (Carver, 1976) and the Singer Eyeball Estimate of Readability or SEER technique (Singer, 1975), this study sought to examine the effectiveness of classroom teachers using these procedures. More specifically, the questions asked were:

1. Are the mean ratings using the Rauding and the SEER technique practically and statistically in reasonable agreement?
2. Do teachers with varying amounts of teaching experience rate passages differently?
3. How do Rauding and SEER ratings compare with ratings based on the Spache or Dale-Chall formula?

Sample

The forty-two teachers enrolled in a fifth-year level university reading course who had completed both sets of ratings were included in the study. Their experience ranged from two to 20 years and the following frequency distribution was used in the analysis:

Experience in Years	*Frequency*
2	5
3	7
4	7
5	4
6 - 10	8
11 - 15	7
16 - 20	4

A 2½ hour lecture on the topic of readability had preceded the rating sessions. The lecture included information about the development of readability formulas, factors of readability, accuracy of the formulas, sources of readability information and special areas of concern (i.e. content areas, examinations). In addition each teacher was required to calculate the readability of a textbook and a tradebook using the Spache formula for primary level materials and the Dale-Chall formula for the intermediate levels and above.

Information about the Rauding and the SEER techniques was presented as part of the above mentioned lecture.

Methodology

The eight passages were taken from Singer's (1975) "unknown paragraphs" but rearranged randomly in two ways (6, 8, 2, 5, 7, 3, 1, 4, and 7, 4, 6, 5, 1, 8, 3, 2) so as to make two forms. An answer sheet with blanks numbered one to eight was provided along with the passages and the criterion passages.

Directions to teachers followed the instructions given by Singer and by Carver in their respective journal articles.

One week the teachers were asked to rate the passages by the SEER technique, the next week by the Rauding index. It was recognized that order effect could not be ruled out by this procedure; however, it was not considered to be an important factor.

An SOL factorial analysis of variance program which could handle unequal replications per cell was used to analyze the data.

Findings

Table 1 presents the mean ratings for each passage (collapsed across years of experience) along with the predetermined readability ratings (Spache or Dale-Chall). The ANOVA results indicated that except for passages #4 (Original paragraph beginning: Several times...) all the remaining passages had statistically significantly different ratings by the two techniques.

The answer to the first question is therefore: No, the Rauding and the SEER technique do not result in similar readability ratings.

A rank order correlation of -0.548 was found between the SEER and Rauding ratings; 0.881 between the SEER and the Spache/Dale-Chall ratings; and -0.762 between the Rauding and the Spache/Dale-Chall ratings when mean ratings for all eight passages were considered.

The second question dealt with differences in ratings by teachers with varying amounts of experience. Table 1 indicates that the ANOVA results on. This conclusion is further supported by the fact that no significant interaction of the eight passages rated. The amount of teaching experience does not appear to affect the accuracy of the ratings in a statistically significant fashieffects were found between the amount of teaching experience and the formula used.

Do the global readability ratings compare favorably with the readability predictions from the Spache/Dale-Chall formulas? Table 1 lists the means (the mean rating of all 42 subjects for each passage) and the Spache or Dale-Chall ratings (provided by Singer, 1975). Even a visual inspection reveals a reasonably close relationship between the SEER and the Spache/Dale-Chall ratings (rho = 0.881). The Rauding ratings appear to result in a reciprocal ordering (rho = -0.762) when compared to formula ratings.

The answer to the last question depends somewhat on which criterion variable one is willing to accept. If the Spache/Dale-Chall ratings are chosen, then the SEER technique is most in agreement. However, the dangers in accepting Dale-Chall ratings have been discussed elsewhere (Froese, 1976).

DISCUSSION AND CONCLUSIONS

Since this study examined only the ratings of eight passages it cannot be

TABLE 1

Analysis of Variance Results, Means, Standard Deviations, and Readability Ratings by Passage

	Passage 1	Passage 2	Passage 3	Passage 4	Passage 5	Passage 6	Passage 7	Passage 8
Formula (df = 1, 6)	114.53**	170.47**	172.66**	0.02	98.55*	39.00**	241.58*	7.89*
Experience (df = 6, 6)	0.98	0.37	3.91	0.73	1.37	1.56	1.57	0.63
Interaction (df = 6, 70)	0.84	1.56	6.04	0.68	0.59	0.24	1.05	0.39
Passage Means (N = 42)								
SEER	4.20	5.62	2.12	4.72	4.75	2.56	2.40	5.09
RAUDING	1.65	2.02	4.80	4.77	2.41	4.10	5.90	4.36
Passage S.D.								
SEER	.902	.888	.277	1.186	1.461	.328	.261	1.592
RAUDING	.754	2.062	1.100	1.987	.853	1.949	1.655	1.139
Spache/Dale-Chall Ratings	4.8	6.4	2.1	4.5	5.9	3.0	1.8	3.8

**significant at .01 level

*significant at .05 level

considered representative of rating all types of materials or levels. However it may be said that the SEER and the Rauding technique do not produce similar ratings on these passages although it should be noted that the range of passages is small (1.8 to 6.4). The reason for the different ratings is not clear but it may be observed that the large, three-grade level intervals on the Rauding scale may not be as appropriate for passages with small variations in difficulty. With the passages rated here it also means that only the bottom half of the Rauding scale is being used.

There appears to be little difference in the ratings of beginning or experienced teachers, and neither technique seems to favor a particular experience level group in a statistically significant manner. However, more guided practice in making readability estimates might well result in better judgements.

If the Spache and Dale-Chall formulas are accepted as the criterion variable, then the SEER ratings are highly correlated (rho = 0.881), and appear to be a reasonable alternative *if several raters are employed.* The ratings of individual passages vary considerably and hence it remains to be shown that reasonable consistency may be maintained for a single rating of a particular passage.

REFERENCES

Baldwin, R.S., & Kaufman, R.K. A concurrent study of the Raygor readability estimate. *Journal of Reading*, 1979, *23*, 149-53.

Bormuth, J.R. Readability: A new approach. *Reading Research Quarterly*, 1966, *2*, 79-132.

Carver, R.P. Measuring prose difficulty using the Rauding Scale. *Reading Research Quarterly*, 1975-76, *11*, 660-684.

Dale, E., & Chall, J.S. A formula for predicting readability. *Educational Research Bulletin*, 1948, *27*, 37-54.

Fry, E.A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-578.

Froese, V. Cloze readability versus the Dale-Chall formula. In B. Schulwitz (Ed.), *Teachers, tangibles, techniques*. Newark: IRA, 1975.

Granowsky, A., & Botel, M. Background for a new syntactic complexity formula. *The Reading Teacher*, 1974, *28*, 31-35.

McLaughlin, G.H. SMOG grading—a new readability formula. *Journal of Reading*, 1969, *12*, 639-646.

Raygor, A.L. The Raygor readability estimate: A quick and easy way to determine difficulty. In P.D. Pearson (Ed.), *Reading: Theory, research and practice*. The 27th Yearbook of the National Reading Conference. Clemson, S.C.: The National Reading Conference, 1977.

Singer, H. The SEER technique: A non-computational procedure for quickly estimating readability level. *Journal of Reading Behavior*, 1975, *7*, 255-267.

Spache, G.D. A new readability formula for primary grade reading materials. *Elementary School Journal*, 1953, *53*, 410-413.

JUDITH WESTPAL IRWIN
Purdue University

TEXT STRUCTURE AND COMPREHENSION: NEW DIRECTIONS FOR READABILITY

Though factors such as sentence length and number of syllables have often been shown to be insufficient causal determinants of readability (see Irwin, 1978), other measurable readability factors have not generally been available. More specifically, recent research indicates that an examination of text-level factors such as coherence and organization might provide readability information not included in word-level or sentence-level analyses. Thus, the purposes of this paper are to summarize current research linking text-level structural characteristics with comprehensibility and to suggest the implications of this research for readability theory and practice.

COHERENCE

Several recent comprehension models have stressed the relationship between coherence and comprehension (Frederiksen, 1977; Hindyard & Olson, 1978; Thorndyke, 1977). Coherence can be loosely defined as what it is that distinguishes a text from a disconnected group of unrelated ideas. Kintsch and van Dijk (1978) have suggested a measure of coherence which, in initial data, correlated .62 with readability, defined as reading time per proposition recalled, and .59 with recall completeness (Kintsch, 1979). Basically, this measure provides a count of the number of instances in which input sets of propositions do not contain an argument already retained in short-term memory. It should be noted that in this study and in Kintsch and Vipond (1977), this measure was a better predictor of readability than a standard readability formula (Flesch, 1948).

Another variable reported in Kintsch (1979) which may provide some measure of coherence is the total number of different arguments contained in the propositional text base of a given text. More different arguments would result in less coherence. Kintsch, Kozminsky, Strey, McKoon, and Keenan (1975) found that the number of word concepts in a passage was significantly related to reading times and recall completeness when the total number of words and propositions in the text was controlled. Similarly, Entin and Klare (1978) report that, in a factor analysis of three correlation matrices of readability variables, a third factor found (after the word-level and sentence-level factors) was the number of different words used in the passage.

Several recent studies have also indicated that coherence may be established by semantic relations other than repeated propositional arguments. Carpenter and Just (1977) found that reading times on a target sentence were significantly shorter when that sentence followed a sentence using a verb that directly entailed a noun phrase in the target sentence. Haberlandt and Bingham (1978) found that reading times were shorter for sentences with verbs that were semantically related, to the previous sentence.

Another possible system for analyzing the coherence of a text is the "cohesion" analysis system presented by Halliday and Hasan (1976). This system includes the grammatical relations of reference (personal, demonstrative, comparative), substitution, ellipsis, and conjuction as well as lexical relations such as exact reiteration and collocation. In a recent study, the present author has attempted to examine the possible relationship between readability, as defined by Kintsch and van Kijk (1978), and cohesion, as defined by Halliday and Hasan (1976), by creating parallel texts with one version containing twice as many cohesive ties as the other. The results indicated that the subjects reading the cohesive version had significantly lower reading-time-per-proposition-recalled ratios and performed better on a prompted recall test than did subjects reading the non-cohesive version.

Moreover, at least four other potential readability factors can be conceptually linked to the idea of coherence. First, the "relational density hypothesis" states that the probability of the recall of any idea is partially a function of the number of other ideas in the text to which that idea is related. It has been supported by the results of two studies (Graesser, 1978). In the first, relational density predicted the recall patterns written by college students immediately after reading. In the second, this hypothesis predicted oral recall patterns collected an hour after reading. If this theory is correct, then it can be predicted that highly coherent texts will be remembered better than less coherent ones.

Secondly, the compactness of subtopic structure has been linked to recall completeness in a study involving sixth-grade students (Aulls, 1975). Topical structure is defined as the "referential connections between statements in a paragraph in terms of each statement's reference to a main topic and subsets of statements to a specific subtopic" (Aulls, 1975, p. 392). When subtopics occur consecutively, they are said to be compact. Clearly, a compact subtopic structure would result in a passage that is more coherent than a passage with a discontinuous subtopic structure.

Thirdly, there is some evidence that the presence of irrelevant information may interfere with comprehension. It could be hypothesized that irrelevant information reduces the coherence of a text. Thorndyke (1979) found that presenting a condensed version of a news story in which irrelevant information was deleted resulted in better recall of the story than did presenting the entire story. Bruning (1970) found that prompted recall of target sentences was greater when those sentences were presented in a relevant context than when they were in an irrelevant context.

Finally, the explicitness of conjunctive concepts has also been linked to comprehensibility (Irwin, 1980b). By definition, when conjunctive relationships between sentences are comprehended, the passage will appear to be more coherent than when they are not comprehended, and recent research indicates that these relationships are more likely to be comprhended when they are stated explicitly. Marshall and Glock (1978) found that subjects from a community college reading a passage in which conditional relations were implicit recalled less of the total passage and fewer of the conditional relations than did similar subjects reading a parallel version in which the conditional relations were explicit. Irwin (1979) found some limited evidence that comprehension of a causal connective is associated with a higher level of recall of the connected ideas, and Irwin (1980a) found that the comprehension of causal connectives may be reduced when they are made implicit.

All of this evidence seems to indicate that one or more measures of coherence would provide useful readability information. This measure would reflect the extent to which sentences are conceptually linked, though it is presently unclear whether this measure should include shared arguments alone, shared arguments along with other semantic relations, a cohesion analysis, a measure of relational density or subtopic compactness, a count of the amount of irrelevant information, an analysis of the number of implicit inter-sentential realtions or some combination of all of these.

ORGANIZATION

Another way of describing text structure is by analyzing larger organizational factors. Indeed, the distinction between coherence and organization is unclear. If one defines coherence as what links sentences to each other, one might describe organization as the abstract structure to which all the sentences are simultaneously linked. Thus, by definition, organization can also be said to contribute to coherence, because all sentences are ultimately linked to each other by their relationship to this abstract structure.

The reader's familiarity with the type of organizational pattern in a text seems to be a potential factor in comprehensibility. In research involving a story grammar describing conventional plot structures, Thorndyke (1977) found that recall was a function of the amount of plot structure in the story and that repeating the same story structure in a second passage facilitated the recall of that passage. Stein and Nezworski (1978) examined the effects of the degradation of their story grammar on college students' free recall and found that recall was "a direct function of the match between the text structure of stories and an ideal story structure" (p. 190). Kintsch (1977) has reported a study which examined the effects of scrambling the order of events in a story. Though it took significantly longer to read the scrambled version, judges of the written summaries could not distinguish between the subjects in the two conditions, indicating that the subjects in the scrambled condition were using the extra time to reorganize the story according to an internalized organizational pattern. Finally, Kintsch and Greene (1978) studied college subjects' recall of a story that corresponded to a predicted story structure and their recall of an Indian story which did not correspond to the predicted story structure. They found that the recall of the "macropropositions", or gist, was severely distorted after five retellings for the Indian story only, therefore concluding that comprehension and recall of stories is facilitated by the availability of an appropriate structural cognitive schema.

If it is true that comprehension is facilitated when the organizational pattern matches an abstract organizational schema already known to the reader, then knowledge of these common organizational schema could make an important contribution to readability determinations. Several possible story structures are available (Kintsch, 1977; Stein & Nezworski, 1978; Thorndyke, 1977), but descriptions of expository patterns are less accessible. Meyer (1978) has suggested that the available schemata for prose organization will probably correspond to the "rhetorical predicates" described by Grimes (1975). In a study involving ninth-grade students, she found that subjects who used the author's rhetorical predicates recalled more information from the passages than the students who did not use these patterns and that students who used this organizational strategy were better able to discriminate between new consistent and intruded information than were the students who did not use this strategy.

Other recent studies have found differential effects for different kinds of organizational patterns. Kulhavy, Shmid, and Walker (1977) found that conceptual and temporal patterns did not result in different levels of recall, but that the subjects reading the concepturally organized passage performed better than those reading the temporally organized passage when conceptual cues were used during recall. Sasson (1971) found that recall of sentence pairs was generally superior when there was a thematic connection, but that temporal connections resulted in better memory for dates. Thus, it might be concluded that the form and content of the desired recall would determine the optimal organizational pattern.

Several studies have compared the comprehensibility of name and attribute forms of organization with conflicting results. Some have found that name organization results in better recall (Divesta, Schultz, & Dangel, 1973; Yekovich & Kulhavy, 1976) and others have found that attribute organization results in better recall (Meyers, Pezdek, & Coulson, 1973; and others). The superiority of attribute patterns seems to disappear, however, when serial cues are not available (Meyers, et al., 1973). Shimmerlik (1978) points out that the materials in these studies varied in terms of such things as content and the availability of other cues and concludes that the superiority of attributive organization seems to depend on the structure of the materials and the specific task involved.

Moreover, the explicit presence of superordinate information has also been linked to comprehensibility. Fishman (1978) and Christie and Schumacher (1978) both report results indicating that an early presentation of superordinate information resulted in higher scores on a recognition test. Greeno and Noreen (1974) conducted a study in which the type of relation among sentences was either categorical, hierarachical, or linear ordering. They found that it took less time to read sentences preceded by superordinate sentences in the hierarchically organized group, but no such effects were found for the other organizational patterns. Gardner and Schumacher (1977) presented a passage in either a superordinate context resulted in higher recall scores only when a less rigorous scoring method was used, concluding that such a context does not aid in the accurate recall of details. Gagne (1969) also varied the type of context for five target facts in a science lesson and found prompted recall to be superior for the subjects in the superordinate-context condition. However, he also found that the subjects in a no-context condition scored higher than any other group and attributed this to contextual interference in the other conditions. Thus, more research is needed to determine the specific situations in which a superordinate context facilitates comprehension.

NEW DIRECTIONS FOR READABILITY

Recent research in text structure and comprehension has several important implications for readability theory and practice. First, text-level factors such as coherence and organization must be considered along with work-level and sentence-level factors when evaluating the readability of a specific set of materials or when attempting to construct parallel experimental passages of equal difficulty. More research is needed in order to determine the most useable measure of coherence with the highest predictive value in a given situation. Extensive text grammars for common expository organizational structure are needed and more research should be conducted on the relative difficulty of various organizational types and on the importance of explicit statements of superordinate concepts.

Second, it is clear that a single-minded application of any factor to all ages and ability levels is likely to be ineffective. Research indicates that such things as the ability to select the main idea (Brown, Smiley, & Lawton, 1977; Brown & Smiley, 1978; Taylor, 1978) and the ability to make intersentential inferences (Bridge, Tierney, & Cera, 1977; Paris & Upton, 1976) develops during the school years. Though sentence length has been shown to be an important variable relating to the reading comprehension of children (Entin & Klare, 1978), it does not seem to have a pronounced effect on mature readers (King & Greeno, 1974; Kintsch & Monk, 1972).

Third, the formula approach must be modified. It is clear that such easily measurable factors as sentence length and number of syllables are not necessary determinants of comprehensibility (Irwin, 1978; Rosenshine, 1969). Moreover, these formulae provide teachers with little useful information about what can be done to increase their students' comprehension and retention. What is needed is an extensive description of all the factors that affect comprehension at various developmental stages. Checklists can be developed which guide teachers of various grade levels to identify salient strengths and weaknesses in their classroom materials (Irwin & Davis, 1980).

Finally, it must also be noted that text structure alone cannot determine readability. All considerations of readability must also include reader-related factors such as prior knowledge, attitude, and motivation (Asher, Hymel, & Wigfield, 1978; Fass & Schumacher, 178; Hittleman, 1973; Klare, 1976). Indeed, research on the relationship of text structure and comprehension must further investigate the interaction of text-structure and reader-related variables if readability determinations are to be optimally effective for specific groups of readers and materials.

REFERENCES

Asher, S. R., Hymel, S., & Wigfield, A. Influence of topic interest on children's reading comprehension. *Journal of Reading Behavior,* 1978, *10,* 35–47.

Aulls, M. W. Expository paragraph properties that influence literal recall. *Journal of Reading Behavior,* 1975, *7,* 391–400.

Bridge, C., Tierney, R., & Cera, M. Inferential operations of children involved in discourse processing. in P. D. Pearson & J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice.* Twenty-seventh Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1978.

Brown, A., Smiley, S. & Lawton, S. *The effects of experience on the selection of suitable retrieval cues for studying from prose passages* (Tech. Rep. No. 53). Urbana-Champaign: Center for the Study of Reading, University of Illinois, 1977.

Brown, A., & Smiley, S. *The development of strategies for studying prose passages* (Tech. Rep. No. 66). Urbana-Champaign: Center for the Study of Reading, University of Illinois, 1978.

Bruning, R. H. Short-term retention of specific factual information in prose contexts of varying organization and relevance. *Journal of Educational Psychology*, 1970, *61*, 186-192.

Carpenter, P. A., & Just, M. A. Reading comprehension as eyes see it. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Christie, D. J., & Schumacher, G. M. Memory for prose: Development of mnemonic strategies and use of higher order relations. *Journal of Reading Behavior*, 1978, *10*, 337-344.

Divesta, F. J., Schultz, C. B., & Dangel, T. R. Passage organization and imposed learning strategies in comprehension and recall of connected discourse. *Memory and Cognition*, 1973, *1*, 471-476.

Entin, E. B., & Klare, G. R. Factor analyses of three correlation matrices of readability variables. *Journal of Reading Behavior*, 1978, *10*, 279-290.

Fass, W., & Schumacher, G. M. Effects of motivation, subject, activity, and readability on the retention of prose materials. *Journal of Educational Psychology*, 1978, *70*, 803-807.

Fishman, A. S. The effects of anaphoric references and noun phrase organizers on paragraph comprehension. *Journal of Reading Behavior*, 1978, *10*, 159-167.

Flesch, R. F. A new readability yardstick. *Journal of Applied Psychology*, 1948, *32*, 221-223.

Fredericksen, C. H. *Inference and the structure of children's discourse*. paper presented for the Symposium on the Development of Discourse Processing Skills, Society for Research in Child Development Meeting, New Orleans, 1977.

Gagne, R. M. Context, isolation, and interference effects on the retention of fact. *Journal of Educationas Psychology*, 1969, *60*, 408-414.

Gardner, E. T., & Schumacher, G. M. Effects of contextual orgainization on prose retention. *Journal of Educational Psychology*, 1977, *69*, 146-151.

Graesser, A. C. How to catch a fish: The memory and representation of common procedures. *Discourse Processes*, 1978, *1*, 72-89.

Greeno, J. G., & Noreen, D. L. Time to read semantically related sentences. *Memory and Cognition*, 1974, *2*, 117-120.

Grimes, J. E. *The thread of discourse*. The Hague, Netherlands: Mouton and Company, 1975.

Haberlandt, K., & Bingham, G. Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 419-425.

Halliday, M., & Hasan, R. *Cohesion in English*. London, England: Longman Group, Ltd., 1976.

Hindyard, A., & Olson, D. R. Memory and inference in the comprehension of oral and written discourse. *Discourse Processes*, 1978, *1*, 91-117.

Hittleman, D. R. Seeking a psycholinguistic definition of readability. *The Reading Teacher*, 1973, *26*, 783-789.

Irwin, J. W. The effects of the interaction of syntactic complexity and the explicitness of connective propositions on fifth-graders' comprehension of prose materials. (Doctoral dissertation, University of Wisconsin-Madison, 1978). *Dissertation Abstracts International*, *39*, 4695-A. (University Microfilms No. 7820627)

Irwin, J. W. Fifth-graders' comprehension of explicit and implicit connective propositions. *Journal of Reading Behavior*, 1979, *11*, 261-271.

Irwin, J. W. The effects of explicitness and clause order on the comprehension of reversible causal relationships. *Reading Research Quarterly*, 1980, *15*, (4), 477-488. (a)

Irwin, J. W. Implicit connectives as a readability factor. *The Reading Teacher*, 1980, *33*, 527-529. (b)

Irwin, J. W., & Davis, C. A. Assessing readability: The checklist approach. *Journal of Reading*, 1980, *24* (2), 124-130.

King, D. R., W. & Greeno, J. G. Invariance of inference times when information was presented in different linguistic formats. *Memory and Cognition*, 1974, *2*, 233-235.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Kintsch, W. On comprehending stories. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Kintsch, W. *On comprehesnion*. Paper presented at the annual meeting of the American Eduational Research Association, San Francisco, 1979.

Kintsch, W., & Greene, E. The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes*, 1978, *1*, 1-13.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. Comprehension and recall of test as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 196-214.

Kintsch, W., & Monk, D. Storage of complex information in memory: Some implications of the speed with which inferences can be made. *Journal of Experimental Psychology*, 1972, *94*, 25-32.

Kintsch, W. & van Kijk, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, *85*, 363-394.

Kintsch, W., & Vipond, D. *Reading comprehension and readability in educational practice and psychological theory*. Paper presented at the Conference on Memory, University of Upsala, June 1977.

Klare, G. R. A second look at the validity of readability formulas. *Journal of Reading Behavior*, 1976, *7*, 129-152.

Kulhavy, R. W., Schmid, R. F., & Walker, C. H. Temporal organization in prose. *American Educational Research Journal*, 1977, *14*, 115-123.

Marshall, N., & Glock, M. D. Comprehension of connected discourse: A study into the relationships between the structure of text and information recalled. *Reading Research Quarterly*, 1978, *1*, 10-56.

Meyer, B. J. F. Organizational patterns in prose and their use in reading. In M. L. Kamil & A. J. Moe (Eds.), Reading research: Studies and applications. Twenty-eighth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1979.

Meyers, J. L., Pezdek, K., & Coulson, D. Effect of prose organization upon free recall. *Journal of Edcuational Psychology*, 1973, *65*, 313-320.

Paris, S. G., & Upton, L. R. Children's memory for inferential relationships in prose. *Child Development*, 1976, *4*, 660-668.

Rosenshine, B. New correlates of readability and listenability. *Reading and Realism*. Proceedings of the Annual meeting of the International Reading Association, 1969, *13*.

Sasson, R. Y. Sematic organization and memory for related sentences. *American Journal of Psychology*, 1971, *84*, 253-267.

Shimmerlik, S. M. Organiztion theory and memory for prose: A review of the literature. *Review of Educational Research*, 1978, *48*, 103-120.

Stein, N. L., & Nezworski, T. The effects of organizational and instructional set on story memory. *Discourse Processes*, 1978, *1*, 177-193.

Taylor, B. M. Children's and adults' recall of general concepts and details after reading. In M. L. Kamil & A. J. Moe (Eds.), Reading research: Studies and applications. Twenty-eighth Yearbook of the National Reading Conference. Clemson, S. C.: The National Reading Conference, 1979.

Thorndyke, P. W. The role of inference in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 437-446.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

Thorndyke, P. W. Knowledge acquisition from newspaper stories. *Discourse Processes*, 1979, *2*, 95-112.

Yekovich, F. R., & Kulhavy, R. W. Structural and contextual effects in the organization of prose. *Journal of Educational Psychology*, 1976, *68*, 626-635.

ROBERT N. KANTOR
ALICE DAVISON
University of Illinois at Urbana-Champaign

CATEGORIES AND STRATEGIES OF ADAPTATION IN CHILDREN'S READING MATERIAL

For the last 30 or 40 years much attention has been paid to formulas that claim to measure readability. The widespread use of such formulas is bound to have an effect on the texts that are produced for children learning to read. In this paper we consider the implications of using readability formulas not just for measuring readability, but also as a guide to producing texts, in this case adapted ones.

The creators of the formulas and others have always warned against using the formulas as guides to production, because the correlation of infrequent vocabulary and long sentences to difficulty in reading does not necessarily imply a causal relationship. Furthermore, it is impossible to use these formulas literally as guides to writing, because provided that one wants to preserve content, the injunction to simplify vocabulary, perhaps by paraphrase, conflicts with the injunction to shorten sentences.

However, because of pressures to produce materials at a specific readability level, writers are often tempted to constrain their writing according to the factors involved in readability formulas. This may affect textbooks for children (and for college students), legal and other documents that must be written in "plain English," and captioning for hearing-impaired people. In our own work, we have examined texts adapted for use by younger readers, and here too have found what we believe to be pervasive effects of readability formulas.

We looked at four texts from SRA Reading Laboratory IIIb (Parker, 1963) designed for students in eighth, ninth, or tenth grade who are reading at levels 5 to 6. We secured the original sources (Glueck, 1962; Mattox, 1961; Melbo, 1941; Parsons, 1913; hereafter referred to as MILK, LIGHT, TREES, and DAYTON, respectively) and did a sentence-by-sentence analysis to determine what information was common to the two versions and what had been changed. All the texts were shortened, average sentence length was lowered, average number of clauses per sentence was reduced, and scores on the Fry and Dale-Chall scales were lowered by zero to five grade levels. We prepared an exhaustive listing of all the changes made in the adaptation. The taxonomy is presented in Davison, Kantor, Hannah, Hermon, Lutz, and Salzillo (1980).

Examination of adapted texts shows that adaptors do not follow readability formulas slavishly. We noted a good deal of conscientious and careful rewriting, but as the examples we will discuss presently show, it seems that vocabulary lists and restrictions on sentence length and passage length are often given primary importance at the expense of factors related to conceptual readability. We will argue in this paper that more attention needs to be given to these factors for which there is to date no objective measurement.

What are these larger considerations that a writer (or adaptor) must be concerned with? The writer should worry first of all about the *overall presentation* of ideas, i.e., the order in which ideas are to be presented and their logical interrelationships. He or she must also be concerned with more *local organization,* i.e., what the topic of discourse is at various points in the text, and where transitions between ideas need to be made. Since the intended reader may not possess requisite *background knowledge* to understand a fact or concept, the writer must sometimes fill in this information for the reader. Finally, the writer must consider matters which influence or shape the reader's *evaluation of information,* e.g., whose attitude is being expressed or whether a statement asserts a fact or is just a supposition.

Our purpose in the following sections is to show what happens to these global considerations when an adaptor operating under a number of different constraints is forced to decide what to change and how to change it. Generally the passage has to be shortened — in our sample anywhere from 32 to 83% — and the end-product must be objectively rated according to at least two readability formulas.

Local Organization

In order to discuss several different aspects of local organization in adaptation, we've chosen a passage from LIGHT, shown in (1) and (2). (1) is the original and (2) is the adapted passage.

(1) LIGHT (original)

(x) In several parts of the world, there are heavy concentrations of these luminous one-celled animals.

(y) Motor launches take visitors into such a lagoon (a) on the southern coast of Puerto Rico where (b) on dark nights there is a dramatic display of luminescence. (c) Curving lines of light fall from the bow (d) as the launch enters the lagoon, and (e) a trail of light is left in the boat's wake. In the lagoon, (f) which has one of the greatest concentrations of bioluminescence in the world, (g) it appears as though a huge floodlight were burning under the launch, and (h) the bow seems to be plowing into a wall of fire.

Sentence structure:

[(y/a) (b)] [(c) (d) (e)] [(f) (g) (h)]

(2) (adaptation)

(a) On the southern coast of Puerto Rico is a lagoon (f) that has one of the greatest amounts of bioluminescence in the world. (b) On dark nights, it creates a very dramatic display. (d/y) As the motor launch takes visitors into the lagoon, (c) curving lines of light fall from the bow. (e) A trail of light is left in the boat's wake. (g) It appears as though a hugh floodlight were burning under the launch, and (h) the bow seems to be plowing into a wall of fire.

Sentence structure:

[(a) (f)] [(b)] [(d/y) (c)] [(e)] [(g) (h)]

This pair of passages illustrates three interdependent phenomena related to local organization that may be affected in the process of adaptation: referential connection, propositional connection, and paragraph organization.

With regard to referential connection, the adaptor has improved on the original in one instance (*such a lagoon* (1y) has no prior discourse referent; the adaptor introduces *lagoon* via an existential construction (2a)), but has allowed the definite description *the motor launch* to appear in (2d/y) without the prior discourse referent that did occur in (1y). Perhaps this occurred because the adaptor's attention was focussed not on the details of referential connection but on the mechanics of restructuring and shortening sentences, as a readability formula would encourage him or her to do.

Note that there are three sentences in the original (disregarding (Ix)) which are restructured into five sentences in the adaptation. The original sentences are more complex, two having three clauses and one having two clauses. Of the adaptation sentences, two consist of one clause and three of two clauses. One effect of splitting up sentences here is that the adaptor is forced to move clause (f) to the beginning of the paragraph, since simply splitting it off would involve loss of propositional connection and lead to incoherence.

Splitting of sentences in adaptation may thus lead to changes in paragraph organization. In this example, the restructuring has the effect of placing all background information in paragraph-initial position, perhaps a more logical ordering, but one that may also involve a loss of stylistic interest.

While in the above example the adaptor seems to have been able to compensate somewhat through ordering for loss of propositional connection, in (3) the splitting forces the reader to guess at the relationship between the propositions.

(3) TREES
- (a) If given a chance before another fire comes, the tree will heal its own wounds by growing new bark over the burned part. (original)
- (b) If given a chance before another fire comes, the tree will heal its own wounds. It will grow new bark over the burned part. (adaptation)

If the reader does not have knowledge about the physiology/biology of trees, or makes an incorrect guess, healing wounds and growing new bark could be seen as two separate processes. This illustrates what may happen if complex sentences are mechanically split up, particularly when connectives which indicate logical relations, i.e., cause, means, sentence, result, are lost in the process.

Background Information

In passages (4) and (5), we find both the deletion and the addition of background information. In (5a) the information that the milk used to flood the rink was surplus has been deleted.

(4) MILK (original)
- (a) In Toronto, a suburban ice-skating rink was flooded with *250 surplus gallons of it* [= skim milk].
- (b) Skaters found it chipped less easily than frozen water.

(5) (adaptation)
- (a) An ice-skating rink was flooded with it [= skim milk].
- (b) Skaters found that *when it froze* it chipped less easily than frozen water.

The adaptor here has deleted the explanation for this otherwise wasteful use of milk. We suspect that the adaptor's motivation was that *surplus* is a difficult or infrequent word. A paraphrase such as "milk that might have spoiled" could have been used, but such a paraphrase would have increased the sentence length and added a subordinate clause.

Comparing (4b) and (5b), we see that the adaptor has added information by inserting a subordinate clause, *when it froze*. This spells out the inference that an adult and likely a child would have no trouble making, i.e., that the milk had to freeze before people skated on it. Obviously the choice of what is deleted and what is added is a matter of intuition on the part of the adaptor, but we are quite surprised that *surplus* was not included and that the fact that the skim milk would freeze was not left to be inferred.

Evaluation of Information

Our final example, comparing original (6a) with adaptation (6b), illustrates what happens to information that a reader might use to evaluate the status of a proposition as a fact or supposition.

(6) TREES

(a) *A railroad freight agent has figured that* it would require at least 40 modern flat cars to haul just the trunk alone. (original)

(b) And at least forty freight cars would be needed to haul away just its trunk. (adaptation)

By deleting the main clause, *A railroad agent has figured that*, the adaptor presents the remaining sentence as an absolute fact. The reader has no basis for evaluation of the statement, except to believe the author unquestioningly. In this case, citing the railway freight agent would lend authority to the proposition, but in other cases, the source of a statement might well have been unreliable. In real life people have to evaluate statements according to their sources, and yet children get little practice in this if source clauses are generally deleted in the material they encounter. We noted many such deletions in our sample. We suspect this information was deleted simply to facilitate the shortening of the sentences.

To conclude, we'd like to offer some suggestions to people who produce reading texts and also to people who buy the texts. Pay less attention to readability formulas and more to content and organization. Several studies (Glazer, 1974; Kaiser, Neils & Floriani, 1975) have already shown that sentence length per se does not necessarily contribute to complexity and hence to difficulty in reading. Psycholinguists and linguists (including ourselves) have argued that sentence complexity is relative to many factors including the discourse context. It is not an absolute value. Those who produce and those who demand texts at guaranteed fixed readability levels may be directing attention to unproductive manipulations of language and away from language which would give children opportunities to develop their conceptual skills.

ACKNOWLEDGEMENTS

The research reported herein was supported in part by the National Institute of Education under Contract No. US-NIE-C-400-76-0116.

No seniority of authorship is intended by the order of authors of this paper. We wish to acknowledge the assistance of Jean Hannah, Gabriella Hermon, Richard Lutz, and Robert Salzillo in carrying out this research.

REFERENCES

Davison, A., Kantor, R., Hannah, J., Hermon, G., Lutz, R., & Salzillo, R. *Limitations of readability formulas in guiding adaptation of texts* (Tech. Rep. No. 162). Urbana: University of Illinois, Center for the Study of Reading, March 1980.

Glazer, S. M. Is sentence length a valid measure of difficulty in readability formulas? *The Reading Teacher*, 1974, *27*(5), 464-468.

Glueck, G. H. About: Grade A. *New York Times Magazine*, February 18, 1962, Section VI, pp. 35-36.

Kaiser, R. A., Neils, C. F., & Floriani, B. P. Syntactic complexity of primary grade reading materials: A preliminary look. *The Reading Teacher*, 1975, *29*(3), 262-266.

Mattox, N. T. Living light on land and sea. *Science Digest*, April 1961, pp. 27-30.

Melbo, I. R. Sequoia National Park. Chapter 11 of I. R. Melbo, *Our country's national parks* (Vol. 2). Indianapolis: Bobbs-Merrill, 1941.

Parker, D. A. *SRA reading laboratory*. Secondary series, IIIb. Chicago: Science Research Associates, 1963.

Parsons, J. Prisoners of the flood. *McClure's Magazine*, 1913, pp. 39-48.

PAPERS PRESENTED AT THE 1979 MEETING OF THE NATIONAL READING CONFERENCE IN SAN ANTONIO

For the convenience of NRC members who may wish to write directly for papers, we are listing all papers presented at the San Antonio conference. We have systematically excluded program chairpersons, discussants, respondents, etc. Also we have tried to make all changes in paper authorship and titles; however it is likely that there still exist some discrepancies.

4.1 Dominic Massato, University of Wisconsin-Madison. Orthography on word perception.

4.2 Ovid Tzeng & Daisy Hung, University of California-Riverside. Orthographic variation in information processing.

4.3 Harry Singer, University of California-Riverside. Orthographies in reading acquisition.

4.4 Richard Venezky, University of Delaware. Orthographic considerations in beginning reading material.

4.5 Marilyn Adams, Bolt, Beranek & Newman. A schema for orthography in reading acquisition.

5.1 Stephen B. Kucer, Indiana University. An international, decision based model of discourse production.

5.2 Nancy L. Shanklin, Indiana University. A model of relationships between two discourse processes.

5.3 Karen Feathers, Indiana University. Language universals in discourse processing.

5.4 Mary W. Hill, Indiana University. A longitudinal study of three and four year olds' discourse processing.

6.1 Emaline B. Henriksen & William Patton, Kent State University. Perceptions of preservice teacher competencies in teaching reading.

6.2 Lenore Ringler, New York University & Ruth S. Meyers, CUNY-Brooklyn College. Teacher interns' conceptualization of reading theory and practice.

6.3 Mary Sue Dillingofski & Ken L. Dulin, University of Wisconsin-Madison. Changes in reading—related attitudes during a one-semester preservice course in secondary reading methods.

6.4 Shirley B. Merlin, James Madison University. A comparison of preservice and inservice teachers' diagnostic reading competencies.

6.5 Gloria T. Blatt, Oakland University. The reading environment in two elementary classrooms, a comparison.

7.1 Bertram Bruce & Cindy Steinberg, Bolt, Beranek & Newman. Interacting plans of characters in stories for children.

7.2 Nancy A. Burris, Incarnate Word College. Role models for readers and student inferred identification.

7.3 Nancy Butler Bowden, University of Houston—Clear Lake City. A study of fourth grade students' reading comprehension measures in short and long passages of an informal reading inventory.

7.4 Michael Johnston & Frank Spencer, University of Arizona. The effects of manipulating the advance organizer on comprehension of abstract material.

7.5 John E. Readence & David Moore, University of Georgia. An integrative analysis of research on structured overviews.

8.1 Gloria A. Schiwetz, University of Houston. A study of creative expression activities and their influence on self esteem.

8.2 Mary Creamer & Jeffrey L. Lorentz, West Georgia College. Interaction of teacher structure, affect cognitive level of questions and group size with student social status as they relate to reading.

8.3 Glen Powell & Robert L. Aaron, University of Georgia. Pupils' on-off task behaviors as a function of teacher race, pupil race, and reading achievement.

8.4 Kathleen Reed, Taunton, MA. Assessing affective responses to reading: non-reactive measures.

8.5 Glen R. Avant, Emory University. The effect of an individualized reading tutorial program on the locus of control and locus of evaluation of disabled readers.

9.0 Walter Pauk, Cornell University. What the college reading program should be doing.

10 Lynne Reder, Carnegie-Mellon University. A discussion of recent research on text comprehension.

11.1 Sharon L. Smith, Indiana University. Organizational schemata and comprehension: a review.

11.2 Robert F. Carey, Indiana University. Contextual constraints as socio-psycholinguistic phenomenon.

11.3 Jerome C. Harste, Indiana University. Comprehension as semantic transaction.

12.1 Randall James Ryder, Kansas State University. Development of variant-predictable letter-sound correspondences in elementary students.

12.2 Connie Juel, University of Texas-Austin & Robert Solso, University of Idaho. Effects of orthographic and phonetic structures on word identification.

12.3 Doris Seeder, Northwestern University. Experience affects children's understanding of before and after.

12.4 David E. Stone, SUNY-Albany. Acquiring knowledge from verbal and pictorial sources.

13.1 William H. Rupley & James F. McNamara, Texas A & M University. Practical significance and statistical models in reading research.

13.2 R. Scott Baldwin, University of Tulsa, John E. Readence, University of Georgia, Thomas W. Bean, California State University & Ernest K. Dishner, Delta State University. An argument for descriptive detail in reading research.

13.3 Prudence Opperman & Eric Brown, New York University. Phonological regression in oral reading.

14. John Downing, University of Victoria. Holmes' sub-strata factor theory revisited.

15. Frank J. Guszak, University of Texas—Austin. Elementary reading programs that work.

16.1 Patricia S. Koskinen & Linda B. Gambrell, University of Maryland. The effects of induced visual imagery upon the reading comprehension of below average readers.

16.2 Cynthia Van Norden Peck, Rutgers University. The effect of imagery and readability on reading comprehension.

16.3 Eugene H. Cramer, University of Illinois-Chicago Circle, & Cheryl Hendress, Project Pegasus, Princeton, Illinois. A comparison of analogical reasoning and reading comprehension ability in middle grade students.

16.4 David Hayes & Patricia L. Anders, University of Arizona. The effect of text-embedded analogy upon comprehension and learning.

17.1 Joyce Pederson, College of St. Thomas. Evaluating the processes and products of the college reading program.

17.2 Dorothy M. Hansen & Sandi H. Amend, Brigham Young University. A required university critical reading evaluation: design, validation and results—a five year report.

17.3 Martha Dillner, University of Houston—Clear Lake. The relationship between performance on the reading and study skills test for medical students and success in medical school.

17.4 Rona F. Flippo, University of South Carolina. The need for comparison studies of college students' reading gains in developmental reading programs using general and specific levels of diagnosis.

18.1 Carol A. Davis & Michael L. Kamil, Purdue University. A pragmatic taxonomy of reading difficulties and symptoms.

18.2 Mark W. Aulls, McGill University. Issues and problems in qualitatively analyzing the reading performance of individuals.

18.3 Margaret MacLean, McGill University. QASOR: a framework for qualitatively analyzing silent and oral reading.

19. Andrew Ortony, University of Illinois. Figurative language: aid or hinderence to comprehension.

20.1 Cindy Steinberg & Bertram Bruce, Bolt Beranek & Newman. A model for the rhetorical structure of narrative and its application to children's literature.

20.2 Bonnie B. Armbruster & Thomas Anderson, University of Illinois. Mapping: representing text structure diagrammatically.

20.3 Robert N. Kantor & Alice Davison, University of Illinois. Categories and strategies of adaptation in children's reading material.

20.4 Dedre Gentner & Andee Rubin, Bolt, Beranek & Newman. A technique to allow practice with high-level aspects of text.

21.1 Carol A. Hodges, University of Illinois. Reading comprehension instruction in the fourth grade: a tale of three methods.

21.2 William Metheny, Michigan State University. The influence of grade and pupil ability levels on teacher conceptions of reading.

21.3 Marie Carducci Bolchazy, Naperville Illinois Unit School District. Teachers' perceptions of students' self-concept and autonomy and indirectness in teaching.

21.4 Luthene A. Chappell, Michigan State University. There is nothing permanent except change (Heraclitus).

22.1 Larry Mikulecky & Jill Edwards Olshavsky, Indiana University. Differences in textbook strategies employed by good and poor university level readers.

22.2 Shirley A. Biggs, University of Pittsburgh. Using the college reader's own strategies to facilitate comprehension.

22.3 Joyce Pederson, College of St. Thomas & Mary A. Moore, University of Wisconsin-LaCrosse. An investigation into the differences between the study techniques of structured overview mapping and student constructed post-questions with college freshmen.

22.4 Ula M. Casale & Brenda Wright Kelley, University of Missouri-Kansas City. Case studies in pass: a problem solving approach to study skills.

23.1 James A. Dinnan, University of Georgia. An evaluation of literacy training for mature adults.

23.2 Molly M. Wilson, University of Georgia. Physical and psychological decrements effecting reading in the aged.

23.3 Albert J. Kingston, University of Georgia. Can reading enhance the lives of the aged?

23.4 Paul Berg, University of South Carolina. Bibliotherapy with the aged.

24. Richard Anderson, The Center for Advanced Studies-Stanford. Schema theory and reading: a progress report.

25. Ronald P. Carver, University of Missouri-Kansas City, versus Ed Coleman, University of Texas. Statistical significance testing, vs replication: merits and demerits of both.

26.1 Carol J. Hopkins, Purdue University. An analysis of individual and institutional levels of productivity for reading journals and conferences.

26.2 Joyce Matheson, University of British Columbia. You're publishing—but who is reading? A study of information seeking behaviors and attitude to information of education practitioners.

26.3 Judith W. Irwin, Purdue University. Text structure and comprehension: readability in the 1980's.

26.4 Victor Froese, University of Manitoba. Rauding and SEERing: judging global readability.

27.1 Joseph Anthony Muia, James Madison University, & Patricia Alexander, Woodstock, Virginia. Reading and the gifted child: what the research tells us!

27.2 Emilie Sullivan, University of Arkansas. Using music to teach reading: state of the art review.

28.1 O. Paul Wielan, University of Georgia. The ability of comparable comprehenders at the 8th, 10th and 12th grade levels to utilize syntactic and semantic redundancy as silent reading strategies.

28.2 Alice Legenza Wescott, University of Kansas & June D. Knafle, University of Illinois-Chicago Circle. Cloze error profiles of German and American adults.

28.3 Roberta C. Johnson & Lea M. McGee, Virginia Polytechnic Institute. Information theory applied to the cloze procedure: effects on comprehension and recall.

29. Robert B. Ruddell, University of California-Berkeley. Literacy use profiles—high and low achievers.

30.1 Magaret MacLean, McGill University. Reading attitude and needs of the aged.

30.2 Richard Allington & Fred Ohnmacht, SUNY-Albany. Task performance as a function of age and reading ability.

30.3 Kathleen M. Ngandu & Bill O'Rourke, University of Nebraska-Lincoln. Older Americans: their reading habits, interests, motivations and attitudes.

31.1 P. David Pearson, University of Illinois. An overview of text comprehension research.

31.2 Don Nix, Watson Research Center. LINKS and the explicit teaching of reading comprehension.

31.3 Wayne Otto & Kay Camperell, University of Wisconsin-Madison. Does text comprehension research offer new insights into the development of marginal gloss strategies.

31.4 Charles W. Peters, Oakland Schools. A content processing model: a new direction for text comprehension research.

32.1 Alden J. Moe, Purdue University. Analyzing text with computers: the state of the technology.

32.2 Franz J. Frederick & Michael L. Kamil, Purdue University. Microcomputers in preservice education of secondary reading teachers.

32.3 Michael L. Kamil & Franz J. Frederick, Purdue University. Solving research problems with microcomputers.

32.4 Ron C. Leslie, New York University. Microcomputers and creative software: research instruments and cognitive analogies.

33.1 Dennis F. Fisher, Human Engineering Laboratory-Aberdeen. Foveal and peripheral visual disfunctioning in disabled readers: implications for compensatory training.

33.2 Yvonne L. Turner, California State Polytechnic University. Right brain vs. left brain in assessing verbal abilities.

33.3 Jimmi Fischer, Gainsville State Correctional Institution, & Judith Cochran, North Texas State University. Identification of right hemispheric remedial readers.

33.4 Judith Cochran, North Texas State University. Remediation for right hemispheric remedial readers.

34.1 Anne D. Forester, Camosun College. Using a process model as a guide to adult reading instruction.

34.2 Sheila Allen, University of Victoria. The use of attitude inventories and needs assessment instruments in teaching the adult student.

34.3 Robert D. Chester, University of British Columbia. Preparing reading instructors for adult basic education.

34.4 David Harrison, Malaspina College. The evaluation of matrials for adult basic reading instruction.

35.1 Roger Pritzkie, Tucson Unified School District & Lisa Howard, Sunnyside Unified School District. The effects of sequencing reading passages on comprehension at the junior high school.

35.2 Marqueritte J. Caldwell, Sunnyside Unified School District, Adela Artola Stewart & Ann Marie Bernazza Haase, University of Arizona. The effects of structural and lexical deletions in cloze procedures on comprehension at the senior high school.

35.3 Curtis Smith, Cochise Junior College. The effects of response types in cloze procedure on comprehension at the junior college.

35.4 Adela Artola Stewart, University of Arizona & Earl Rankin, University of Kentucky. Issues in material development for experimenting and teaching with cloze procedure.

36.1 Patricia Baggett, University of Denver. Structurally equivalent stories in movie and text: processing and recall.

36.2 Joel R. Levin, University of Wisconsin-Madison. The role of pictures and imagery strategies in prose comprehension.

36.3 Mary Beth Marr, SUNY-Albany. Picture and text sequences: processing differences.

36.4 S. Jay Samuels, University of Minnesota. The role of pictures in sight word learning: a review and critique.

37.1 Bruce L. Brown & Dorothy M. Hansen, Brigham Young University. New evidence for an old issue: an empirical demonstration of rapid reading.

37.2 James V. Hoffman & Sharon O'Neil, University of Texas-Austin. An investigation into the internal and external rate flexibility of proficient readers in relation to the difficulty level of the material.

37.3 Marilyn G. Eanet, Rhode Island College & Jane Warren Meeks, Old Dominion University. Flexibility of reading rate: a fact or an artifact of measurement?

37.4 Mark Condon, University of Louisville. Internal flexibility of oral and silent reading rates in the rauding mode.

38.1 Rebecca Kirschenman, Ohio State University. Retelling: the state of the art.

38.2 Lowell Eberwein & Jean Pival, University of Kentucky. Effects of repatterning a story on oral reading miscues and comprehension.

38.3 Josephine Spivack Goldsmith, Rutgers University. Examining miscue assumptions: towards a new taxonomy.

38.4 Jerome A. Niles, Virginia Polytechnic Institute. The effects of selected teacher prompting strategies on oral reading performance.

38.5 Donna L. Wiseman, University of Missouri-Columbia. A psycholinguistic description of the reading and writing behavior of a selected group of five year old children.

39 Rand Spiro, Yale University. New perspectives on schema theory and reading comprehension.

40.1 Eileen B. Entin & George R. Klare, Ohio University. Cloze tests of comprehension on ambiguous passages.

40.2 Michael C. McKenna & John W. Miller, Wichita State University. The effects of age, reading ability, and distractor type on maze performance.

40.3 Patricia S. Gingrich, New York University. Clozing the gap: an investigation of conjunctions.

40.4 Susan Thomas, Fayette County Schools, & Earl F. Rankin, University of Kentucky. The effect of prior knowledge upon the operation of contextual constraints in cloze passages.

41.1 Karl Koenke & Sheryl Benson, University of Illinois. A validation of the revised Artley-Hardin inventory of teacher knowledge of reading.

41.2 Emery P. Bliesmer & George S. Johnson, Pennsylvania State University. Spelling at high school and adult levels: exploratory research.

41.3 Thomas W. Bean & Thomas C. Potter, California State University-Northridge, & Charles H. Clark, Western Illinois University. Selected semantic features of ESL materials and their effect on bilingual students' comprehension.

41.4 Eunice L. Schmidt, Seattle Pacific University. Linguistic awareness of young readers in Denmark and the United States.

42.1 Jane L. Davidson & DeWayne Triplett, Northern Illinois University. Thematic convergence in reconstructed passages in mapping activities.

42.2 David Caverly, Indiana University. Rehearsal as a retention strategy in reading: developing a theoretical base for teaching study systems.

42.3 Barbara Taylor & Sandra Berkowitz, University of Minnesota. Facilitating children's comprehension of content area material.

42.4 Margaret Policastro, Pamela Adelman, Bernice Goldberg, Doris Seeder & Elizabeth Sulzby, Northwestern University. The effects of three instructional techniques on the comprehension of a well-formed story.

43.1 Martha Rapp Haggard, Northern Illinois University. Vocabulary acquisition during elementary and post-elementary years.

43.2 Joe Peterson, Kansas State University. Relationships between meaningfulness and ease of acquisition of sight vocabulary with first and fifth grade subjects.

43.3 Robert L. Aaron & Glen Powell, University of Georgia. The relationship of intelligence to rate for learning reading vocabulary.

43.4 Michael F. Graves, University of Minnesota, Heidi Hammond, West St. Paul Public Schools, & Judith A. Boettcher, University of Minnesota. A validated procedure for teaching prefixes and its effect on students' ability to assign meaning to novel words.

44. Eric Brown, New York University. The relevance of neurological factors in the reading process.

45.1 Sandra Bulke, Michigan State University. You have to be a part of it to know what I mean! Methodological evolution of a naturalistic study.

45.2 John Stansell & Carol Robeck, Texas A & M University. The effect of professional training on theoretical orientation to reading.

45.3 Diane E. DeFord, Southern Illinois University. The effect of years of experience on changes in theoretical orientation due to instruction on the college level.

46.1 Charles H. Clark, Western Illinois University, & Judith W. Irwin, Purdue University. The development of macro-level textual components in the comprehension of organized and unorganized prose.

46.2 James L. Laffey, James Madison University, & Donna Kelley, Dayton, VA. The influence of repeated reading on reading achievement and attitudes toward reading.

46.3 James W. Cunningham, University of North Carolina-Chapel Hill, & Patricia M. Cunningham, Wake Forest University. Further investigating the "print-to-meaning" hypothesis.

47.1 Ann J. Pace, University of North Carolina-Greensboro. Metacomprehension and school-aged children's reading: an overview.

47.2 Victoria Chou Hare & Cynthia A. Pulliam, University of Illinois-Chicago Circle. Adult readers' metacognitive awareness of reading strategies.

47.3 Linda K. Crafton, Indiana University. Experiential background: the reader's inexorable contribution to the reading process.

47.4 John H. Mize, Alamance County Schools, NC. The relationship between reading performance and awareness of comprehension of textbook passages.

48.1 Judith A. Langer, New York University. Pre-reading language and concept organizers: a schema theory application.

48.2 Phyllis Hawley & Allan Neilsen, University of Calgary. Effects of prior knowledge and linguistic connectives on comprehension of causal relationships in written discourse.

48.3 Mariam Jean Dreher, Richard P. Salley, & Harry Singer, University of California-Riverside, & Virginia L. Berry, Rialto Unified School District. Recall of direct vs. inferred instantiations in children.

48.4 Nancy Marshall, Montgomery County Public Schools, MD. The effects of discourse type upon comprehension.

49.1 Kathleen Lyon & Mary Beth Marr, SUNY-Albany. Passage independency of the informal reading inventories.

49.2 Edward Fry, Rutgers University. A passage dependent test to provide a different view of reading comprehension.

49.3 Alton L. Raygor, University of Minnesota. The Minnesota reading assessment: a new tool for community college, post-secondary, and adult programs.

50.1 Mary Doyle, North Texas State University. The role children's interests have taken in shaping the reading program: research overview.

50.2 Ruth Berrier, Hurst-Euless-Bedford ISD, TX. Developing a reading interest inventory: primary and intermediate.

50.3 M. Jean Greenlaw, North Texas State University. Using case study and interest inventories: a research report.

50.4 Sylvia Carter, University of Georgia. Teacher's and pupil's perception of sex roles as they relate to reading interests.

51.1 Beverly S. Morrison, University of Wisconsin-Madison. Reading/learning skills: how relevant are they? Reactions of secondary level teachers in seven content areas.

51.2 Eunice N. Askov, Pennsylvania State University, Karlyn Kamm, University of Wisconsin-Madison, & Roger Klumb, University of Maine. Study skills mastery: comparisons between teachers and students.

51.3 Vincent P. Orlando, Oberlin College. A comparison of notetaking strategies while studying from text.

51.4 Nancy Kendall Snouffer & Linda Thistlethwaite, University of Western Illinois. Use of the structured overview in two content areas.

51.5 Richard F. Barron & Robert M. Schwartz, Oakland University. The effect of student constructed graphic post-organizers upon concept differentiation in the teaching of reading.

52.1 Mary Ann Siebert, University of Houston. Portrayal of the teacher in picture books.

52.2 Eleanore S. Tyson, University of Houston. Portrayal of the teacher in primary books.

52.3 Rosemary Oliphant Ingham, University of Houston. Portrayal of the teacher in intermediate novels.

52.4 Marian Perry, University of Houston. Portrayal of the teacher in adolescent novels.

53.1 Sandra White & Pauline Witte, University of Wisconsin-Madison. Improving comprehension through repetition... repetition... repetition.

53.2 M. Jane Greenewald, University of Wisconsin-LaCrosse & Anne E. Wolf, University of Wisconsin-Eau Claire. The status and use of reading and other language arts in junior and senior high content-area instruction.

53.3 Terry Lovelace, University of Southwestern Louisiana & Sylvia Duchamp, University of Southern Louisiana. Effects of individual differences on secondary students' performance on selected learning tasks.

53.4 Mary Lorton, Aquinas College. Effect of languaging curriculum on college freshmen.

53.5 Ken L. Dulin, University of Wisconsin-Madison, & M. Jane Greenewald, University of Wisconsin-LaCrosse. Male and female college-age readers' judgments of the gender appropriateness of informal colloquial, and substandard dialogue associated with male and female literary characters.

54.1 Gerald G. Duffy, Michigan State University. Direct instruction in reading: an analysis of the instructional practices of twenty-three elementary school teachers.

54.2 Taffy E. Raphael, University of Minnesota, Peter Winograd & P. David Pearson, University of Illinois. Strategy perceptions and actual performance in question answering tasks.

54.3 Jane Hansen, University of Minnesota. Changing children's questioning environments: its effects on comprehension.

54.4 Christine Gordon, University of Minnesota. Effects of instruction in metacomprehension and inferencing on children's comprehension abilities.

54.5 Ellen Sylvester, University of Minnesota, Taffy Raphael, & P. David Pearson, University of Illinois. Prior knowledge and reading comprehension: changing states of schemata.

55.1 Elizabeth Sulzby, Northwestern University. The development of metalinguistic understandings about words.

55.2 Bernadette Pudis & Elizabeth Sulzby, Northwestern University. The development of children's understanding of the linguistic unit "word".

55.3 Nelly M. Hecker, Furman University & Bob W. Jerrolds, University of Georgia. Space test scores as compared to cloze and multiple-choice.

55.4 Charles Temple, University of Houston-Victoria. Learning to spell in other languages.

55.5 R. Darrell Morris, University of Virginia. The beginning reader's concept of word and its relationships to phonetic segmentation ability.

5.6 M. Jane Kita, University of Virginia. Children's concepts of reading and writing.

55.7 Jean Wallace Gillet, University of Virginia. Classroom implications of word concept research.

56.1 Patricia L. Anders & Judy Nicols Mitchell, University of Arizona. A cross-grade thematic evaluation of nine social studies texts: relational text-patterns, supporting activities, and concept development.

56.2 William Earl Smith, Indiana University. The psychological reality of a story grammar and its relationship to reading comprehension.

56.3 Randolph K. Cirilo, University of Texas-Austin. Some effects of story structure during comprehension.

56.4 V. Frank Stone, Western Illinois University. The effect of textual cohesion on the comprehension of connected discourse.

56.5 Paula J. Gaus, University of Arizona. The effects of three types of metaphor on sixth grade students' reading comprehension.

56.6 R. Timothy Rush, Purdue University. Predicting first-grade reading achievement with case grammar aspects of spoken discourse: an application of Kintsch's system of propositional analysis.

57.1 Harry Bowman & Paul Jones, Memphis State University. U.S. navy recruit phonic abilities and referral to remedial instruction.

57.2 June D. Knafle, University of Illinois-Chicago Circle & Mary Ann Geissal, Northeastern Illinois University. Variations in attitudes toward children's oral reading errors.

57.3 Rebecca J. Palmer, Wayne H. Slater & Michael F. Graves, University of Minnesota. The effect of passage difficulty on good and poor readers' use of authors' schemata in written recall protocols.

57.4 Myrtle Mayer, SUNY-Albany. Developmental analysis of good and poor readers in recognition of word targets in three orienting conditions.

57.5 Pauline F. Moley & John E. George, University of Missouri-Kansas City. Effects of applying whole to part and part to whole reading instructional approaches to disabled readers with identifiable cognitive learning styles.

57.6 Timothy Shanahan, University of Delaware. Three theories of reading disability: a literature review.

57.7 Estelle L. Fryburg, Manhattan College. The relationship between reading failure and learning disabilities among delinquent youth.

58.1 Ernest T. Goetz, University of Illinois. The use of context by good and poor readers in a cloze task.

58.2 James W. Pichert, Vanderbilt University. Sensitivity to importance as a predictor of reading comprehension.

58.3 Ralph E. Reynolds, University of Illinois. The effect of inserted questions on the intensity and duration of text processing.

58.4 Diane L. Shallert, University of Arizona. A comparison of content and structural contexts in learning from expository prose.

58.5 Robert M. Schwartz, Oakland University. Levels of context utilization in the reading process.

59. Carl Frederiksen, McGill University. Inference in children's comprehension of texts.

60.1 Bertha Wilhelm Kingore, North Texas State University. An ethnographic investigation of the influence of the home environment on the developing reading skills of first graders.

60.2 Barbara Thompson, North Texas State University. Parents' perceptions of their roles in children's development.

60.3 Ruth J. Kurth, North Texas State University. Research on parent-child communication.

60.4 Jane Boyce Seibert & Barbara Mathews, Dallas Independent School District, Carol Colvin, N. Texas State University. Emerging trends in parent-school communication.

61.1 Betty H. Yarborough & Roger A. Johnson, Old Dominion University. A longitudinal study of sex differences in reading among elementary-school pupils.

61.2 Carol K. Weber, C.U.N.Y.—York College, & Lenore H. Ringler, New York University. Motivational and knowledge influences on text comprehension.

61.3 Bonnie Longnion & Jesus Garcia, Texas A & M University. Publishers respond to sexism in basal and supplementary readers: a change in the nature of stories.

61.4 Richard F. Abrahamson & Rosemary Oliphant Ingham, University of Houston. A structural approach to children's expressed reading preferences.

62.1 Dennis Guiser, Michigan State University. Influence of reading instruction on the development of writing among first graders.

62.2 Marilyn M. Fairbanks & Mary K. Elliott, West Virginia University. Providing reading aids and evaluating the effect of selected background factors on student achievement in a composition course for high risk freshmen.

62.3 Christine W. Lucas, McGill University. Systematic reading-writing for student and teacher competency in the language of instruction.

63.1 Anita V. Kak, Northwestern Illinois University. Cognitive style: its role in the development of the reading process.

63.2 Charles W. Peters, Oakland Schools. Measuring reading performance in four cognate areas through the utilization of metacognitive scale.

63.3 Vicky Lynn Miller, University of Texas-Austin. An investigation of the cognitive processing of children with good and poor reading comprehension.

63.4 David M. Wark & J. Michael Bennett, University of Minnesota-St. Paul. Impact of test anxiety on reading comprehension: test of a model.

64.1 Priscilla A. Drum, University of California-Santa Barbara. Relationships between text and question characteristics and correct answers in comprehension tests.

64.2 Robert J. Tierney & James Mosenthal, University of Arizona. Reader, text and texture.

64.3 James Flood, Boston University. Paraphrasing ambiguous written discourse: comprehension and production.

64.4 Carol N. Dixon, University of California-Santa Barbara. Children's listening and reading recall strategies related to text type.

64.5 Taffy Raphael, Peter Winograd & P. David Pearson, University of Illinois. Relationships between children's perception and performance in strategies for answering questions.

64.6 Diane Lapp, Boston University. Directing questioning strategies through discourse analysis application.

65.1 Patricia H. Duncan, Virginia Commonwealth University. Family education in reading: a school-based model.

65.2 Frederic Linder, Virginia Commonwealth University. Evaluation models: parent education in reading.

65.3 Nancy Boraks, Virginia Commonwealth University. Coping strategies of parents with minimal reading skills: ethographic study.

65.4 J. Lee Pemberton, Virginia State Department of Education. The role of the state education agency in parent involvement.

66.1 Joseph L. Vaughan, Jr. & Diane Shallert, University of Arizona. Identification of prerequisite schemata; implications for reading research.

66.2 Thomas H. Estes & Wayne L. Shebilske, University of Virginia. Comprehension as cooperative structuring of text.

66.3 Susan Thomas & Connie Bridge, University of Kentucky. The relationship between subjects' cloze scores and their ability to employ macrostructure operations in the generation of summaries.

66.4 James Mosenthal & Robert J. Tierney, University of Arizona. Macrostructures and macro-operations during discourse comprehension.